Studies on the structure and function of the influenza virus ribonucleoprotein and polymerase complex.

Thesis

UNRESTRICTED

THESIS

submitted by

Klaus Klumpp, Dipl.Biol.

for the degree of

Doctor of Philosophy of the Open University, Milton Keynes

in the field of biology

(submitted 1 November 1996)

Studies on the structure and function of the influenza virus ribonucleoprotein and polymerase complex

sponsoring establishment:
GlaxoWellcome
Research & Development
U.K.

collaborating establishment:
EMBL Grenoble Outstation
France

Date of submission: 1 November 1996
Date of award: 21 January 1997

ABSTRACT

Genomic RNPs were isolated from influenza A virus particles and the RNA conformation in the native structure was studied by chemical and enzymatical RNA modification analysis on complete RNPs and on RNPs depleted of polymerase protein. The results suggest that NP binding to the genomic RNA does not involve the Watson-Crick positions of the bases and the major part of the genomic RNA was found to be in a single-stranded conformation, consistent with the idea that NP functions as a single-stranded RNA binding protein, that facilitates templated RNA synthesis by the polymerase. The polymerase, on the other hand, was found to form a ternary complex with both conserved vRNA ends, but in the absence of the polymerase the RNA ends were completely single-stranded despite a partial, inverted complementarity. The implications of the absence of a stable panhandle RNA secondary structure on possible replicative mechanisms are discussed.

RNP activity assays were employed to determine basic kinetic constants of the transcription reaction catalyzed by RNPs and to analyse the mechanism of inhibition by 2'-deoxy-2'-fluororibonucleotides, which were found to be incorporated into RNA products by the influenza virus polymerase thereby preventing efficient elongation of the transcripts. Transcription reactions in the presence of ATP analogs and ATPase assays revealed a possible requirement for ATP β-γ bond hydrolysis during influenza virus transcription.

Finally, a number of different protein expression systems were tested for the production and purification of the PA subunit of the influenza virus polymerase complex for structural and functional studies. In both procaryotic and eucaryotic systems the PA protein yield was very low and possible reasons for this lack of expression efficiency are discussed. Affinity purified PA preparations from bacterial cultures and semipurified fractions from eucaryotic systems showed that an ATPase activity consistently copurified with PA. Also, preliminary evidence for sequence-specific RNA binding was obtained from *in vitro* translation experiments.

Acknowledgements

The work for this thesis has been performed between January 1994 and October 1996 in the groups of Dr. Martin Ford and Dr. Rob Ruigrok at Wellcome Research Laboratories, Beckenham, U.K. and EMBL Grenoble Outstation, France respectively.

I would like to express my sincere gratitude to my supervisors Drs. Martin Ford and Rob Ruigrok for giving me the opportunity to work in their laboratories, for their ideas and interest in the ongoing project and their availability for discussions of all kinds of problems.

Many thanks to Drs. John Skehel, John McCauley, Graham Darby and Rob Ruigrok for reading the thesis manuscript and accepting to be my examiners.

Many thanks also to Dr. Margaret Tisdale for all support, helpful discussions and for keeping an eye on the university programme regulations. I am indebted to Dr. Florence Baudin for introducing me to the techniques of RNA structure analysis, and I also want to thank Dr. Stephen Cusack for his support, his interest in this work and for some stimulating discussions.

Special thanks to Dr. Kjeld Larsen for helping me with the HPLC, Dr. Filippo Volpe in the baculovirus laboratory and Christine Vincent for her extraordinary pioneering work on *Pichia* systems and encouraging me to follow her into the yeast phase. I would also like to thank all the other people in the labs at Beckenham and Grenoble for the nice atmosphere, discussions and generous sharing of knowledge.

This work was funded by GlaxoWellcome Research and Development, U.K.

Abbreviations

A	adenine
aa	amino acid
AMPCPP	α,β-methyleneadenosine 5'-triphosphate
AMPPCP	β,γ-methyleneadenosine 5'-triphosphate
AMPPNP	5'-adenylylimidodiphosphate
ApG	adenylyl(3'-5')guanosine
ATP	adenosine 5'-triphosphate
β–SH	β-mercaptoethanol
bp	base pair
C	cytosine
CD	circular dichroism
cDNA	complementary DNA
CMP	cytidine 5'-monophosphate
cpm	counts per minute
cRNA	complementary RNA
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
ddATP	2',3'-dideoxyadenosine 5'-triphosphate
ddCTP	2',3'-dideoxycytidine 5'-triphosphate
ddGTP	2',3'-dideoxyguanosine 5'-triphosphate
ddTTP	2',3'-dideoxythymidine 5'-triphosphate
dGTP	2',3'-deoxyguanosine 5'-triphosphate
DMS	dimethyl sulfate
DNA	deoxyribonucleic acid
DOC	deoxycholic acid
DTT	dithiothreitol
dTTP	2'-deoxythymidine 5'-triphosphate
EDTA	ethylenediaminetetraacetic acid
EtBr	ethidium bromide
f	fempto (10^{-15})
FdCTP	2'-deoxy-2'-fluorocytidine 5'-triphosphate
FdGTP	2'-deoxy-2'-fluoroguanosine 5'-triphosphate
g	gramme
g	earth's gravitational field (9.81 ms^{-2})
G	guanine
GTP	guanosine 5'-triphosphate
h	hour
HA	hemagglutinin
HIV	human immunodeficiency virus
HPLC	high pressure liquid chromatography
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobases (10^3 bases)
kD	kilodalton (10^3 Dalton)
K_m	Michaelis-Menten constant
M	molar (1 mole per litre)
m	metre
m	milli (10^{-3})
MBP	maltose binding protein
MDCK cells	Madin Darby canine kidney cells
mg	milligramme (10^{-3} gramme)
min	minute
ml	millilitre (10^{-3} litre)
mM	millimolar (10^{-3} mole per litre)
mRNA	messenger RNA

Abbreviations

n	nano (10^{-9})
NA	neuraminidase
nm	nanometre (10-9 metre)
NP	nucleoprotein
nt	nucleotide
NTP	nucleotide 5'-triphosphate
OD_{600}	optic density at 600 nm
p	pico (10^{-12})
p.i.	post infection
p.i.	post induction
PA	influenza virus polymerase subunit PA
PA-wt	wild-type PA
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
pCp	cytidine(3', 5')bisphosphate
PCR	polymerase chain reaction
pfu	plaque forming units
pI	isoelectric point
pmol	picomole (10^{-12} mole)
pol	polymerase
RNA	ribonucleic acid
RNP	ribonucleoprotein
RNP-pol	ribonucleoprotein devoid of polymerase
rpm	revolutions per minute
RT	room temperature
s	second
S	sedimentation coefficient (Svedberg unit, 10^{-13} seconds)
SDS	sodium dodecyl sulfate
TCA	trichloroacetic acid
tRNA	transfer RNA
U	uracil
UMP	uridine 5'-monophosphate
UTP	uridine 5'-triphosphate
UV	ultraviolet
vRNA	viral, genomic RNA
WHO	world health organization
μ	micro (10^{-6})
μg	microgramme (10^{-6} gramme)
μl	microlitre (10^{-6} litre)
μM	micromolar (10^{-6} mole per litre)

Table of Contents

1. Thesis Introduction

1.1. Influenza, the disease

Influenza is an infectious, acute respiratory disease caused by the influenza virus, and it is characterized by an abrupt onset of symptoms like headache, high fever, chills and significant myalgias. Fever usually declines at day 2-3 after the onset of the symptoms, but severe cough and general malaise can persist for 1-2 weeks. The infection affects and partially destroys the ciliated epithelium of the respiratory tract, the regeneration of which can take up to 1 month to be completed. Severe complications comprise secondary bacterial infections and viral pneumonia with mortality rates up to 42% (Tashiro et al., 1987; Murphy and Webster, 1990). Influenza virus infections can be particularly dangerous for people belonging to certain high risk groups like newborn children, people above 65 years of age, pregnant women or people with chronic pulmonary or heart diseases. Considering the number of affected people worldwide, every year, influenza constitutes one of the major infectious diseases. The WHO coordinates a surveillance program to monitor the epidemiology and the emergence of new virus strains [http://www.who.ch/]. Influenza virus is unique among the respiratory tract viruses in that it continually undergoes significant antigenic variation. The current vaccines, that are based on inactivated viruses, have to be modified every year (Kilbourne, 1987). However, they are of limited efficacy especially during the regularly recurring pandemic outbreaks, that are characterized by a significant increase in the number and the severity of cases. The pandemics are caused by newly emerging influenza virus subtype strains due to genetic reassortment of virus strains. The incorporation of envelope proteins from animal virus sources into a viral genome optimized for infection in humans can lead to highly virulent, new virus strains. The detailed study and characterization of the influenza virus with molecular biology and biochemical techniques will provide information on the structure of the virus, its genetic variation and its mode of gene expression and replication, and it will eventually lead to new approaches for the control of influenza viruses. At the same time the study of the mechanisms involved in the virus life cycle will help to better understand the biology of many related viruses and elucidate general principles in eucaryotic molecular and cellular biology and immunology (e.g., the concept of mRNAs as intermediates in gene expression was originally established in the *E.coli*/T2 bacteriophage system (Brenner et al., 1961)).

1.2. Influenza Virus Structure

Influenza virus is a membrane-enveloped, negative-strand RNA virus with a genome size of about 14 kb. It shares these qualities with members of 3 other virus families, that all comprise important human pathogens, the Rhabdoviridae (e.g. vesicular stomatitis virus, rabies virus), Filoviridae (Marburg and Ebola viruses) and the Paramyxoviridae, (e.g. parainfluenza viruses, mumps virus, measles virus, respiratory syncytial virus). Because of some very unique features, influenza viruses have been placed into their own family, the Orthomyxoviridae. The most characteristic properties are a segmented genome structure with 10 proteins expressed from 8 different RNA molecules, a high rate of genetic recombination (mainly through reassortment) and genome replication in the nucleus of infected cells. The influenza viruses have been classified into 3 types (A, B and C) according to serological cross-reactivities of internal antigens, but they also differ in epidemiology, evolutionary patterns and molecular fine-structure. In particular, influenza C virus genomes contain only 7 RNA segments and encode only one outer envelope protein. Most of the work, including the present study, has been done on influenza A viruses, that show the most dramatic variability, but in general influenza B viruses have been found to follow very similar molecular mechanisms. The influenza virions usually are 90-120 nm in diameter (figure 1A), but also filamentous forms with lengths up to several micrometers have been observed with human isolates after a single round in cell culture (Choppin et al., 1960; Chu et al., 1949). The membrane envelope contains 3 different proteins:

- HA (haemagglutinin) forms about 400 trimeric spikes on the outside of a characteristic virus particle, well visible on electron micrographs. HA is responsible for the attachment of the virus to sialic acid containing target cell membranes and for the fusion of viral and cellular membranes during virus entry.
- NA (neuraminidase) is present in about 100 mushroom-shaped tetramers. Its enzymatic activity hydrolyzes terminal sialic acid residues and is required for the passage of the virus through mucus layers in the respiratory tract and for the release of virions during the budding process (White, 1974).
-M2 protein forms about 3-17 tetrameric proton channels in the virus envelope and is possibly involved in genome release during cell entry (Zebedee and Lamb, 1988).

In the interior of the virus particle, the M1 protein forms a shell surrounding the viral genome and attaching it to the envelope (Fujiyoshi et al., 1994; Ruigrok et al., 1989). This layer may consist of approximately 3000 copies of the M1 protein, as determined by biochemical protein analysis and cryoelectron microscopy (Compans et al., 1970; Fujiyoshi et al., 1994; Skehel and Schild, 1971). Biophysical mass determination and comparison of complete and bromelain treated virus particles support a virus structural model with a M1 protein network consisting of only 1200 M1 monomers (Cusack et al., 1985; Ruigrok et al., 1984).

The influenza virus genome consists of single-stranded RNA segments (vRNAs), which are packed by nucleoprotein (NP) into structurally distinct ribonucleoprotein particles (RNPs). The RNPs are associated with a heterotrimeric, RNA-dependent RNA polymerase composed of the subunits PA, PB1 and PB2 (Lamb, 1989). The vRNA terminal sequences of all genome segments are highly conserved despite a relatively high mutation rate (2-7 * 10^{-3} per nucleotide and year) in the rest of the sequence (Smith and Palese, 1989). The 13 nucleotides at the 5' end are completely conserved among all influenza A virus strains, and the 12 nucleotides at the 3' end show only a single base difference in certain RNA segments. The vRNA ends also show a partial, inverted complementarity over a region of 15-16 nucleotides, which means that they can theoretically base-pair with each other and form a so-called panhandle structure (Desselberger et al., 1980; Robertson, 1979; Skehel and Hay, 1978; Stoeckle et al., 1987). Support for the possible formation of such a RNA secondary structure *in vivo* came from psoralen crosslinking studies (Hsu et al., 1987). The RNPs are the functional units for viral transcription and replication, that take place in the nucleus of infected cells (Herz et al., 1981; Jackson et al., 1982). The RNP structure, as apparent from electron microscopy, seems to be mainly determined by the NP polymers (Pons et al., 1969; Ruigrok and Baudin, 1995). The RNPs appear as strands with loops at one or both ends and with different lengths corresponding to the sizes of the RNAs they contain (Compans et al., 1972). Figure 1B shows an exemplary electromicrograph of RNP obtained with a carbon shadowing technique. The RNP tertiary structure is formed by NP inducing a coiled structure of the RNA. The coil itself is folded back on itself and arranged into a large, helical structure with the dimensions of about 10-20 nm diameter and 50-130 nm length (Compans et al., 1972; Pons et al., 1969; Compans et al., 1972; Pons et al., 1969; Jennings et al., 1983). The polymerase trimer has been located at, or very close to, the end of each RNP by immunogold labelling (Murti et al., 1988). At present it is not known how the RNPs are sorted and packaged during virus assembly and how they are arranged inside virus particles.

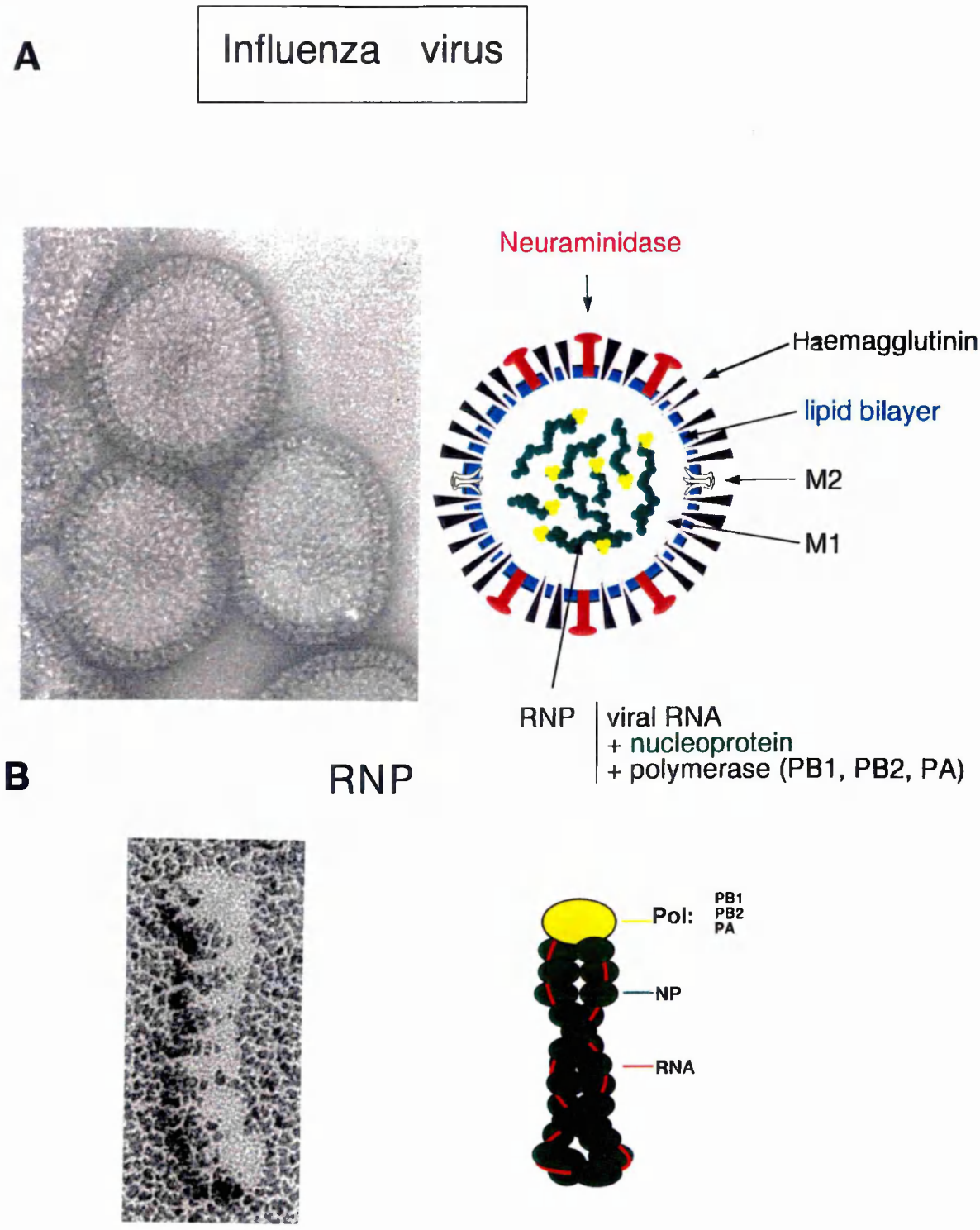


Figure 1: Influenza virus structure. Electron micrographs (left, courtesy of Dr. Annie Barge) and diagram (right) of influenza virus (A) and ribonucleoprotein, RNP (B). The viral membrane contains three viral transmembrane proteins, hemagglutinin, neuraminidase and M2.The M1 protein forms a shell surrounding the genomic RNP particles. RNP consists of a genomic RNA molecule, nucleoprotein (NP) and a trimeric polymerase complex.

1.3. Influenza virus Life Cycle and Host Cell Interactions

Viruses are unable to generate metabolic energy or to synthesize proteins by themselves. They have to enter host cells to replicate, which often results in an exploitation of host cell molecules at the expense of the cell, one of the basic causes of viral disease. Influenza virus infection eventually leads to cell death by apoptosis via the PKR (double-stranded RNA activated protein kinase) pathway (Hinshaw et al., 1994; Takizawa et al., 1995; Takizawa et al., 1993). Also other viruses, e.g. HIV 1, Epstein-Barr virus and Sindbis virus, induce apoptosis, which is considered as a host defense mechanism to eliminate the pathogens. Accordingly, viruses are known to produce proteins that counteract or delay this host cell suicide response until they have produced sufficient amounts of progeny virus (Chou and Roizman, 1992; Clem et al., 1991; Rao et al., 1992).

The virus life cycle can be divided into three major phases: cell entry, genome replication and release of progeny virus (figure 2). Important aspects of each phase in influenza virus infection are briefly summarized below.

1.3.1. Virus Entry

The first event in viral infection of a host cell is binding of the virus to the cell surface. The influenza virus haemagglutinin (HA) is responsible for viral attachment to cellular sialyloglycolipids or sialyloglycoproteins. The virus particles are then internalized via clathrin coated pits and they fuse with the cell membrane in the late endosome compartment, when the pH has dropped below a certain threshold level. The pH drop leads to a conformational change in the HA protein causing the exposure of a hydrophobic peptide, that promotes fusion between the two membranes (Murphy and Webster, 1990; Skehel et al., 1982). The low pH conditions in the endosome environment are probably also important for the uncoating of the genome from the M protein shell, possibly due to the acidification of the virus interior through the M2 proton channel (Martin and Helenius, 1991a; Zhirnov, 1990). Finally, virus uptake and transfer into the late endosomes has the advantage of providing transport for the RNPs through possible obstructions in the cytoplasmic compartment delivering them directly into the central perinuclear area (Marsh and Helenius, 1989). The M1-protein-free RNPs are then actively transported into the cell nucleus via the nuclear pore complex (Martin and Helenius, 1991b).

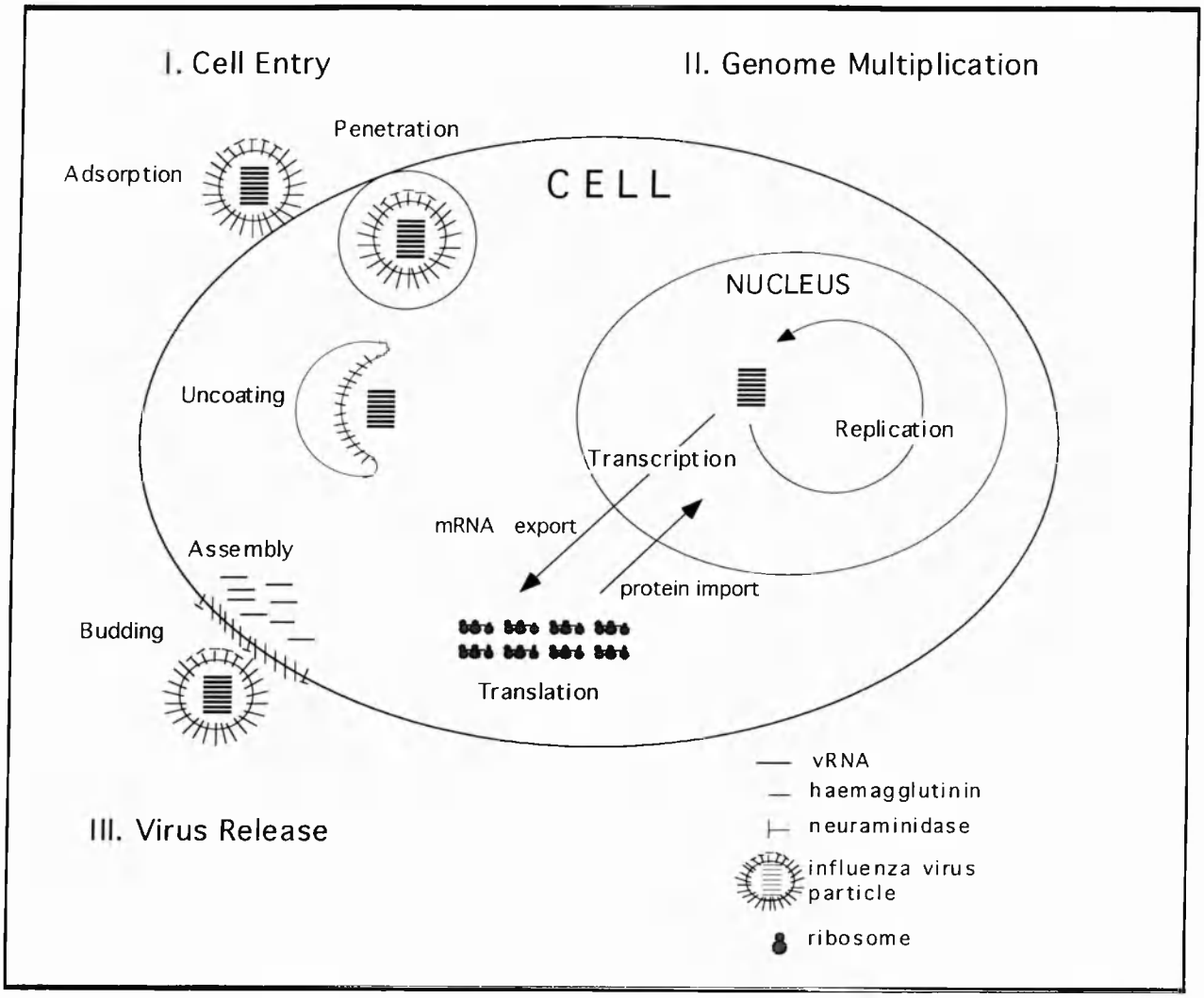


Figure 2: Cycle of influenza virus infection schematically depicting the series of events that lead to virus multiplication in susceptible cells from the onset of adsorption to final budding and release of progeny virus

2.3.2. Gene Expression (transcription and replication)

Influenza virus transcription and replication take place in the nucleus of infected cells in close association with the nuclear matrix (Jackson et al., 1982; Lopez-Turiso et al., 1990). Whereas the genomic RNA of (+) strand viruses can directly function as mRNA in infected cells, (-) strand viruses carry specific polymerases, that produce mRNAs by transcription of their genomic vRNAs. Because influenza viruses contain their specific RNA-dependent RNA polymerase on the RNPs, it was initially surprising to find that the virus was dependent on functional, cellular RNA polymerase II for transcription (Rott et al., 1965; Rott and Scholtissek, 1970; Lamb and Choppin, 1977). The influenza virus polymerase harbours a unique endonuclease activity, that cleaves RNA polymerase II derived hnRNAs or mRNAs of the cell at certain positions 10-15 nucleotides from their

cap structures. These capped oligoribonucleotides are then used as primers to initiate viral transcription from the second base of the vRNA 3' end (Caton and Robertson, 1980; Dhar et al., 1980; Krug et al., 1979). The cap snatching strategy of the influenza virus not only provides the necessary mRNA structures for efficient translation, but also strongly interferes with cellular gene expression and is probably one of the causes for the shut-off of host protein synthesis in infected cells (Lazarowitz et al., 1971; Skehel, 1972). Transcription terminates at an oligoU stretch 17-22 nucleotides from the 5' end of the vRNA, where the polymerase adds a polyA tail of 140-160 nucleotides to the mRNA by a presumed stuttering mechanism (Hay et al., 1977; Krug et al., 1989; Plotch and Krug, 1977). The first step in the replication of the influenza virus genome is the switch from mRNA synthesis to the synthesis of cRNA, the full-length copy of genomic vRNA. The cRNAs then serve as templates for vRNA production. During replication the RNAs are synthesized without cap groups and polyA tails and the transcription termination signal is ignored (Hay et al., 1982; Skehel and Hay, 1978; Young and Content, 1971).

These processes are subject to a tight regulation, that results in the synthesis of different amounts of proteins at different times of infection, which can be divided into two distinct phases. In the early phase (until 2-3 h *p.i.*) mRNA, cRNA and vRNA synthesis occur in parallel and control of gene expression is mainly achieved by the selective production of different amounts of vRNA for different segments. In this way the polymerase subunits, NP and NS1 proteins are preferentially synthesized. Later in infection mRNA and cRNA synthesis are drastically reduced, whereas the production of vRNA and proteins continues from persisting cRNA and mRNA templates (Krug et al., 1989; Shapiro et al., 1987). The mechanisms involved in these crucial regulatory processes are only poorly understood. It has, however, been established that newly produced NP protein early in infection is required by the polymerase in order to pass through the termination signal and produce full-length cRNAs and vRNAs *in vitro* (Beaton and Krug, 1984; Shapiro and Krug, 1988).

Generally, virus multiplication involves a lot of specific interactions with the host cell. Influenza virus is unique among RNA viruses in producing two different proteins from single mRNAs by alternative splicing. The splicing is performed by cellular spliceosomes and is tightly regulated to produce distinct amounts of each protein. Influenza virus gene products (NS1 and polymerase) are involved in the control of the splicing process in favour of viral gene expression. NS1 protein inhibits splicing by specifically capturing an essential component of the spliceosome, the U6snRNA (Lu et al., 1994; Qiu et al., 1995). Additionally, NS1 protein interferes with the nuclear export of polyA containing mRNAs, presumably favouring the export of virus specific messages (Fortes et al., 1994; Qiu and Krug, 1994). The virus also has to cope with the consequences of interferon action and PKR activation in infected cells (see above). Enhanced translation of viral mRNAs and a virtually complete shut-down of host protein

expression is presumably achieved in several ways. The virus very efficiently blocks the activation of the kinase PKR by activating a cellular PKR-inhibitor, called p58 (Lee et al., 1990; Lee et al., 1992). Additionally, NS1 protein seems to block PKR activation by binding to double-stranded RNA (Lu et al., 1995). Active PKR phosphorylates and thereby sequesters in an inactive form the translation elongation factor eF2-α. Not only does influenza virus block this kinase, but viral mRNAs seem to be capable of specifically recruiting eF2-α, when this factor becomes limiting due to extensive phosphorylation (Feigenblum and Schneider, 1993). The same authors observed a significant dephosphorylation of the cap-binding initiation factor eIF-4E during influenza virus infection, which leads to a significant downregulation of cap-dependent transcription initiation. In addition, eIF-4E is the most limiting factor relative to other initiation factors and constitutes a major target for translational control by a large number of effectors, e.g. growth factors (Sonenberg, 1993). Influenza virus mRNAs apparently overcome this initiation block, because they require extremely low amounts of this initiation factor due to the short length and the absence of secondary structure in the 5' non-translated region (Feigenblum and Schneider, 1993). The conserved 5' end of influenza virus mRNAs has been shown to be sufficient to confer translational advantage to chimeric RNAs. Therefore, a binding protein, specific for the conserved 5' end, may be involved in the influenza virus specific upregulation of translation (Garfinkel and Katze, 1993). These examples emphasize how the influenza virus exploits cellular mechanisms for its own replication and can even take advantage of cell defense mechanisms.

2.3.3. Release of progeny virus

The genomic vRNAs are presumably packaged by NP into RNP structures during RNA synthesis, which occurs in the nucleus. Then they are actively and unidirectionally transported into the cytoplasm to be incorporated into progeny virus particles. This transport is dependent on M1 protein entering the nucleus, but the mechanism of the RNP export process is unknown (Martin and Helenius, 1991a; Whittaker et al., 1996). Influenza virus assembly and budding take place at the plasma membrane of infected cells, or the apical membrane in the case of epithelial cells (Kingsbury, 1990). Relatively large amounts of the proton channel forming M2 protein are synthesized in infected cells. This protein is presumably needed to equilibrate the pH in the trans Golgi network (TGN), which prevents the premature conformational change of some HA subtypes to the low pH form (Ciampor et al., 1992; Sakaguchi et al., 1996). During virus assembly M2 protein is apparently excluded from virus particles (Zebedee and Lamb, 1988). The enzymatic activity of the NA protein cleaves off sialic residues from neighbouring HA molecules and from any cell surface molecules in the vicinity. This enables the virus to leave the cell membrane and prevents virus particles from aggregating (Kingsbury, 1990; Palese et al., 1974).

2.4. Antiviral Therapy

The currently available vaccines based on inactivated viruses are usually quite effective in reducing the frequency and the severity of the influenza disease, and annual vaccinations are recommended for people belonging to high risk groups (Murphy and Webster, 1990; Ruben, 1987). However, the vaccines are not completely protective against influenza virus infection. Maximal efficacy of 80% is only seen with homologous virus challenge and new vaccines have to be produced annually in order to keep up with the prevalent virus strains, that undergo continuous changes by antigenic drift. In the worst case, major changes of the surface antigens, haemagglutinin and neuraminidase, can occur by recombination and significantly diminish or nullify the protection conferred by the vaccines. Antiviral agents are needed for the treatment of acute infections and they should ideally be targeted at virus-specific, essential steps of the replicative cycle without affecting normal host cell metabolism. The influenza virus infection cycle provides a number of potential targets for the development of antiviral agents. Presently, the only licensed anti-influenza drugs are amantadine and rimantadine, which are of limited utility, because they very rapidly generate transmissible, resistant influenza virus mutants and they show adverse renal and neurological side effects (Hayden and Hay, 1992). Both compounds block the ion channel activity of the M2 protein and inhibit virus uncoating and assembly events (Bukrinskaya et al., 1982; Hay et al., 1985; Martin and Helenius, 1991a; Pinto et al., 1992). Resistant viruses have mutations in the transmembrane region of the M2 protein (Hayden and Hay, 1992; Houck et al., 1995). A number of alternative compounds have also been investigated as potential anti-influenza drugs (figure 3).

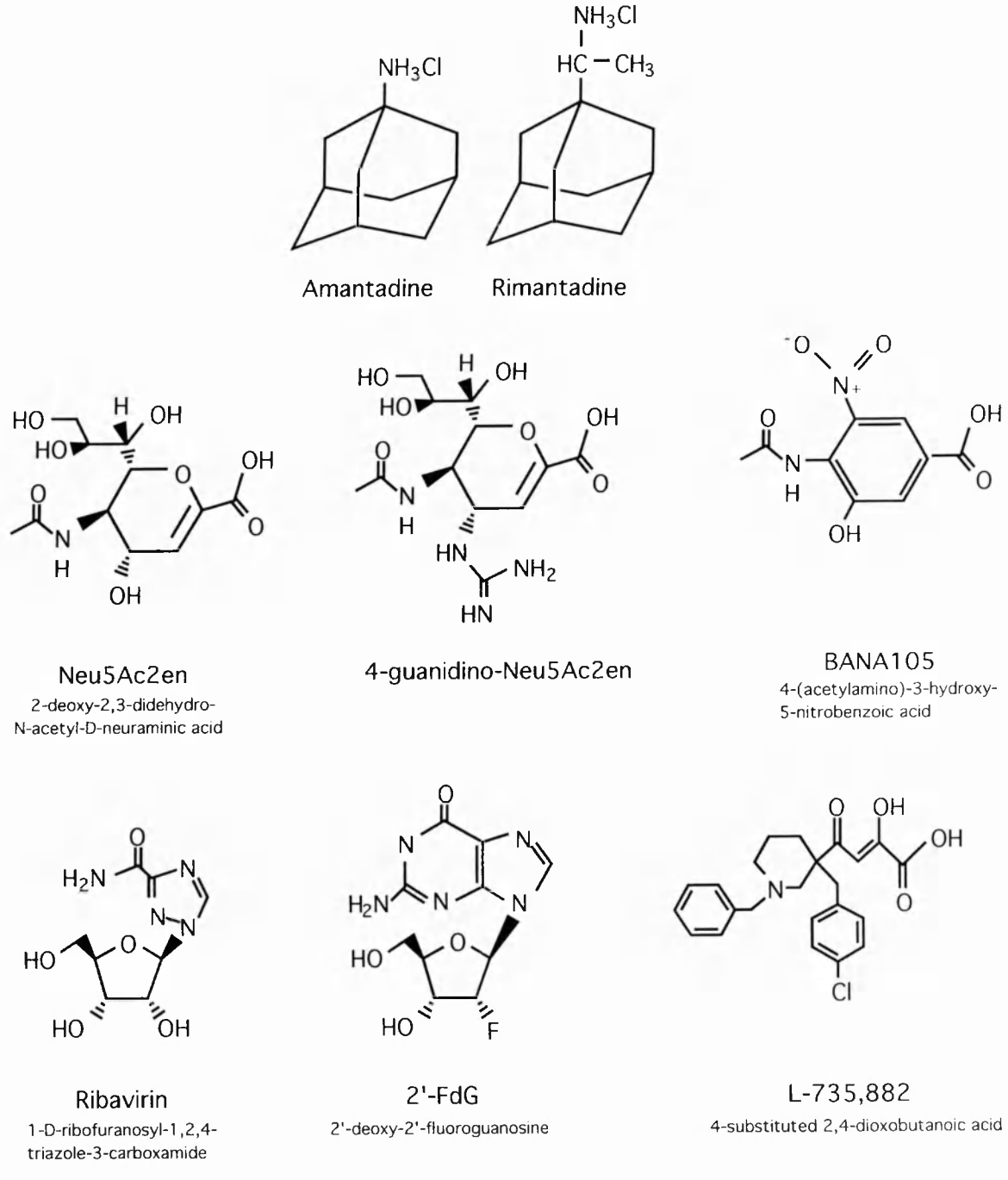


Figure 3: Molecular structure diagrams of compounds, that have shown considerable anti-influenza virus activity *in vitro* and *in vivo*

The sialidase activity of the NA protein constitutes another virus-specific target located on the envelope of the virus particles. Knowledge of the NA crystal structure led to a significant improvement in the efficacy of sialidase inhibitors, which are transition state analogs of the sialidase reaction. The neuraminic acid derivatives show significant NA-

inhibiting activity *in vitro* and efficiently prevent virus shedding and pyrexia *in vivo* (Schulman and Palese, 1975; von Itzstein et al., 1993b). However, in spite of the conservative nature of the NA active site, differences in drug sensitivity were observed among different NA subtypes, and the compounds were completely ineffective against some highly virulent, avian influenza A virus strains (Gubareva et al., 1995). Other problems comprise rapid excretion, very short-term protection from virus challenge and poor oral bioavailability (Colman, 1994). However, randomized, double-blind, placebo-controlled trials with healthy, young adults showed that the NA inhibitor 4-guanidino-Neu5Ac2en (GG167) was safe and effective for prophylaxis and early treatment of experimental human influenza with a H1N1 influenza virus strain (Hayden et al. 1996). Alternative compounds mimicking the structure of neuraminic acid derivatives are currently being developed, based on the structure and interaction of the inhibitors with NA, to address these problems (Jedrzejas et al., 1995). The search for drugs interfering with the interaction of haemagglutinin and sialic acid was unsuccesful until now, presumably due to relatively weak interaction between these two molecules and the lack of charge-charge interactions in the sialic acid binding site of HA (von Itzstein et al., 1993a).

The majority of the presently licensed antiviral agents are nucleoside analogs, that interfere with virus-specific nucleic acid synthesis (Hirsch and Kaplan, 1990). The influenza virus RNA-dependent RNA polymerase also possesses a number of unique activities, that constitute ideal drug targets. Ribavirin was the first nucleoside analog, that inhibited influenza virus replication, but its antiviral spectrum is larger and includes several other RNA viruses like respiratory syncytial virus (RSV) and Lassa fever virus (McCormick et al., 1986; McIntosh et al., 1984). The mechanism of ribavirin action is complex, the drug interferes with cellular GTP pool sizes and a number of adverse side effects are known, which limit the utiliy of this compound.

2'-deoxy-2'-fluororibosides have shown significant inhibitory activity against influenza A and B viruses in standard and human tracheal cell cultures, and they very efficiently decreased virus titres in infected mouse lungs. Especially the guanosine analog was significanty more efficient than both ribavirin and amantadine in the mouse model (Tisdale et al., 1993). Because nucleoside analog compounds are targeted to the highly conserved viral polymerase protein, the emergence of resistant viruses is likely to be significantly delayed compared to amantadine.

The endonuclease activity of the polymerase complex is even more specific to the influenza virus, than the RNA-dependent RNA polymerase module itself. Several laboratories have recently been trying to establish endonuclease activity assays to screen for influenza virus-specific inhibitors. Dioxobutanoic acid derivatives have turned out to specifically inhibit the mRNA cleavage reaction, but not transcription initiation or elongation by the influenza virus polymerase protein (Tomassini et al., 1994). Although the first series of compounds could not completely inhibit virus replication in MDCK

cells, the results demonstrate that the influenza virus polymerase can be a specific target for antiviral agents. Further structural and biochemical characterization of this protein complex will considerably facilitate the development of such compounds.

2.5. Objectives

The goal of this study was to better understand the structure-function relationship of the influenza A virus polymerase in the context of the transcriptionally active unit of the virus, the RNP. The influenza virus polymerase is the central protein in the virus infection cycle and harbours most of the activities necessary for the transcription and replication of the viral RNA genome. The polymerase is composed of three virally encoded protein subunits PB1, PB2 and PA, which form a heterotrimeric complex *in vitro* and in *vivo* (Akkina et al., 1987; Detjen et al., 1987; Hagen et al., 1994; Jones et al., 1986; Krug et al., 1989). The polymerase strongly associates with the genomic RNA, and this interaction can be stable enough for purification on CsCl gradients (Honda et al., 1988). The exact polymerase binding site on the RNA, however, has only been poorly defined (Honda et al., 1987; Murti et al., 1988). Considering such a strong interaction on virus-derived RNPs, major changes in protein-RNA contacts are likely to occur in order to enable polymerase progression and template copying during transcription and replication. The study of the promoter structure for transcription has been considerably hampered by the lack of commendable sources of purified or recombinant polymerase protein. Purification protocols of polymerase from viruses or infected cells either do not remove all endogenous viral RNA or largely denature the protein (Honda et al., 1990; Seong and Brownlee, 1992; Szewczyk et al., 1988). Recombinant expression systems on the other hand usually produce polymerase protein only in minute amounts or are deficient in complex formation (Hagen et al., 1995; Kobayashi et al., 1992) Another important factor, that could not be considered in most polymerase-RNA binding studies performed so far, is the substantial influence of the nucleoprotein (NP) on the RNA structure (Baudin et al., 1994). We have therefore decided to work with intact genomes purified from the virus and studied the polymerase-RNA interaction in the presence of NP by chemical modification analysis. The results of this work are presented in chapter I. These studies have provided more detailed information on the RNP structure and the interactions of both polymerase and NP with vRNA.

Recombinant expression systems for polymerase subunits and efficient reconstitution systems will be essential to find out more about the polymerase organization itself and to obtain structural information of subunits and active sites. It is still far from understood which subunits make contacts with the RNA genome and how they do so. The polymerase also catalyzes a variety of different reactions which are both essential and highly specific for influenza viruses. In most cases, the important regions on the proteins have not been mapped yet. Especially there is still not much evidence

concerning the function of PA in the viral replication cycle. Studies with temperature sensitive viruses suggested that PA function was mainly required during replication, but possibly not during primary transcription (Krug et al., 1975; Palese et al., 1977). We have therefore cloned the PA subunit gene into a variety of expression vectors in order to find a system that might provide us with sufficient amounts of protein for functional or even structural studies. Chapter II compares the results obtained with different expression systems and discusses their potential use to determine PA function.

As outlined above, the unique activities of the influenza virus polymerase constitute ideal targets for antiviral chemotherapy. Chapter III describes work on *in vitro* transcription assays in order to determine the mechanism of transcription inhibition by 2'-fluororibonucleotides. These compounds may become valuable tools to study early transcriptional processes and specific polymerase activities *in vitro*.

Chapter I

RNA modification analysis of native RNPs

I.1. Introduction

Influenza virus RNPs, composed of RNA, nucleoprotein (NP) and polymerase, can be considered as the minimal units to perform viral RNA synthesis. The selective separation of NP from polymerase-RNA complexes *in vitro* leads to a significant loss of transcription activity and mainly affects elongation (Honda et al., 1988). This suggests that the polymerase complex contains the enzymatic activities needed for transcription initiation, whereas NP has mainly a structural role, presumably in presenting the template RNA to the polymerase in a conformation optimal for transcription elongation. *In vivo*, the expression and replication of virus-like RNAs requires the polymerase and NP as the minimal set of proteins (de la Luna et al., 1993; Huang et al., 1990; Kimura et al., 1992; Mena et al., 1994). The RNP structure (see introduction, page 4, figure 1) cannot be functionally reconstituted by simply mixing purified proteins with RNA, although the resulting structures resemble RNPs on electron micrographs (Honda et al., 1988; Kingsbury et al., 1987; Krug et al., 1989). RNA modification analysis was employed to study the RNA conformation in the context of the intact RNP structure and to obtain information on the protein-RNA interactions within this structure.

I.2. RNA modification analysis: a method to study the conformation of RNA molecules in solution

The methodology to study RNA conformation in solution is based on the availability of a variety of specific chemical and enzymatic probes, that allow the analysis of single nucleotides along a RNA molecule (reviewed by Ehresmann et al., 1987). These probes indicate if specific nucleotide positions are either freely accessible from the solvent or protected by a RNA secondary structure or a bound protein. This method is based on the chemical sequencing reactions developed in the laboratory of Walter Gilbert (Peattie and Gilbert, 1980) and has been used extensively in the study of protein-RNA complexes, e.g. in spliceosomes (Jandrositz and Guthrie, 1995), ribosomes (Holmberg et al., 1992; Moazed et al., 1988) or regulatory mRNP complexes (Li et al., 1995). Figure I.1 illustrates the principle of the method. In the first step, the RNPs are enzymatically cleaved or chemically modified at a statistical and low level (less than one cut or modification per molecule). After stopping the reaction, the RNAs are purified by phenol extraction and hybridized to a radioactively labelled DNA oligonucleotide complementary to a chosen sequence on the RNA.

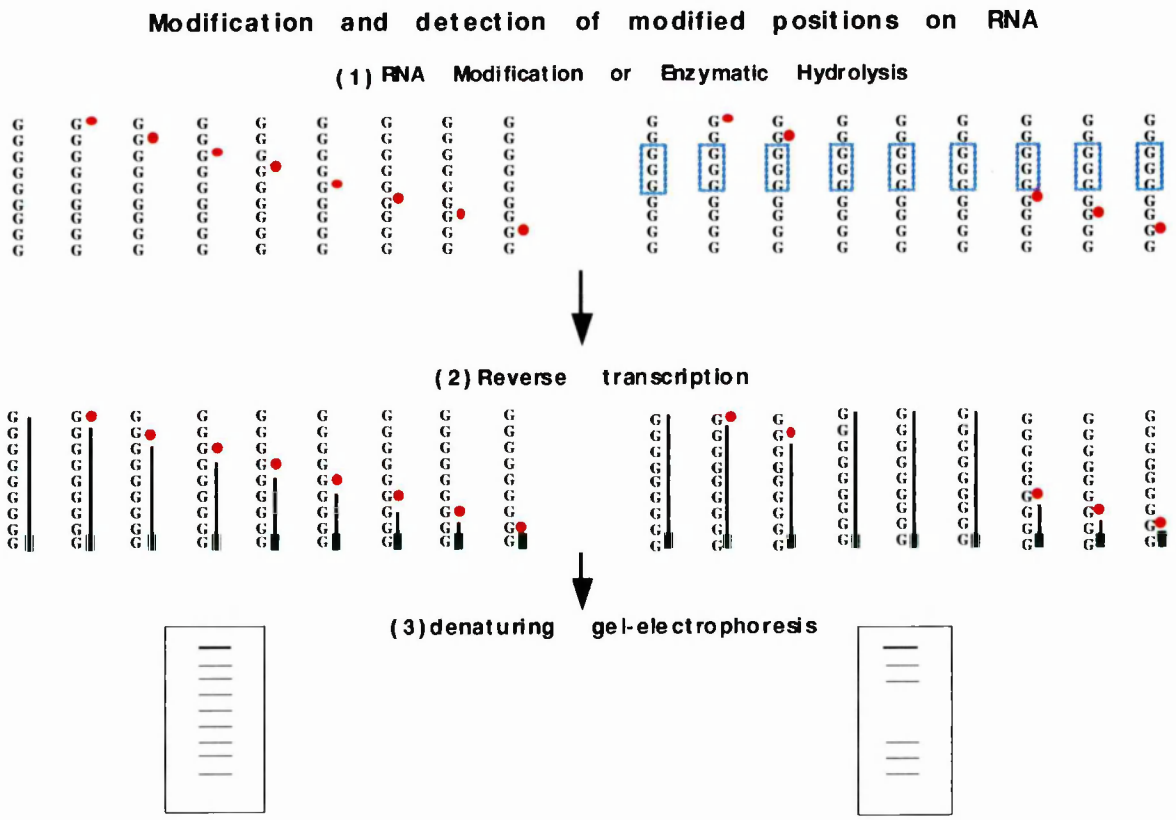


Figure I.1: Schematic representation of the general principle of RNA modification analysis and the primer extension method for the detection of modified bases. (GGGGGGGG) stands for a random RNA sequence, (●) indicates modified bases, (▯) a protected region (binding protein or RNA secondary structure), (▬) represents a labelled primer DNA and (▬—) an extended primer.

The cleavage sites or the modified bases are then identified in a primer extension reaction with reverse transcriptase. The elongation during reverse transcription is prematurely terminated at cleavage sites or one nucleotide before a base, that has been modified at a Watson-Crick position. The cDNA fragments synthesized by this procedure are labeled at their 5' ends and can be visualized by autoradiography after denaturing gel electrophoresis (urea-PAGE). A sequence-specific ladder of bands will be produced on such gels, when single-stranded, naked RNA is analyzed (figure I.1, left). However, RNA binding proteins or RNA secondary structure elements can protect bases from being modified, which results in the absence of a band at the corresponding place on the gel (figure I.1, right). Control reactions are performed in parallel by incubating the RNAs in the absence of modifying agents. In this case the primer extension reaction primarily produces full-length transcripts. Additional bands in the control reactions can be caused by nicks in the RNA or unspecific stops of the reverse transcriptase.

The reverse transcription reaction only resolves RNA sequences upstream of the chosen primer binding site. An alternative approach has to be used to identify modified base positions at the extreme 3' ends of RNA molecules. In the present study, the modified

influenza virus RNPs were phenol-extracted, and the resulting vRNAs 3' end-labelled with T4 RNA Ligase and radioactive cytidine-3'-5'-bisphosphate (pCp). The labelled RNAs could then be specifically cleaved at the modified positions by a further chemical treatment. This approach requires base-specific chemicals modifying the bases in such a way, that the RNA can be chemically cleaved after the modification reaction. Dimethylsulfate (DMS) is at present the only chemical, that sequence-specifically interacts with Watson-Crick positions of bases and cytidine residues can afterwards be detected by chemical RNA cleavage.

I.2.1. Chemistry of modification probes

Figure I.2 shows a Watson-Crick base-pairing scheme illustrating the sites of base reactivity with the probes, that have been used in this work.

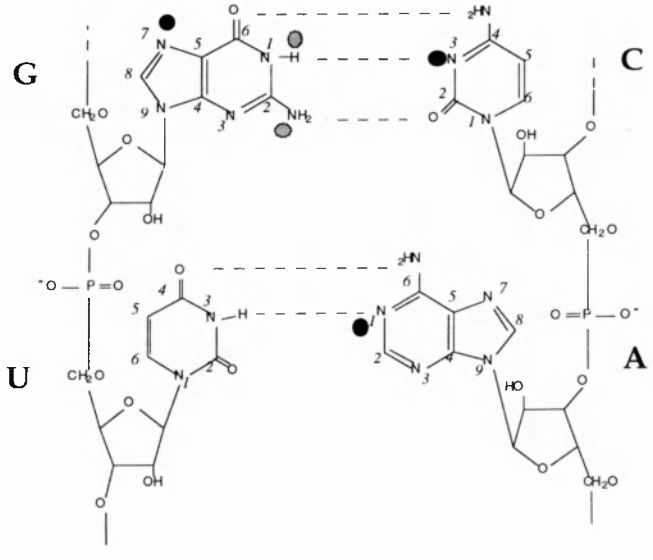


Figure I.2: Watson-Crick base-pairing scheme illustrating the interference of base-pairing interactions with the sites of modification by DMS (●) and kethoxal (◎).

Dimethylsulfate (DMS):

At neutral pH DMS methylates positions N7 of guanines (G), N1 of adenosines (A) and N3 of cytosines (C). The methylated C and A can be detected by reverse transcription. The RNA can also be specifically cleaved at DMS modified C residues by a sequential treatment with hydrazine and aniline (Ehresmann et al., 1987; Peattie and Gilbert, 1980).

β-ethoxy-α -ketobutyraldehyde (Kethoxal)

Kethoxal interacts with the Watson-Crick positions N1 and N2 of G residues. The reaction forms a new ring structure, which is unstable in basic conditions. Borate ions are added after the reaction to stabilize the product.

RNAse T1

RNAse T1 forms specific hydrogen bonds with N1 and O6 of G residues before hydrolyzing the RNA via the formation of an intermediate guanosine 2'-3'-cyclic phosphate. The hydrolysis generates fragments with a 3' phosphate and can be used as an indicator for unpaired G.

I.3. Modification analysis of native influenza virus RNPs

I.3.1. Purification of RNPs

Influenza A/PR/8/34 virus RNPs were purified from egg-grown virus according to a protocol, that is currently used in similar forms in several laboratories (see chapter IV B1.2. as adapted from Honda et al., 1988). Briefly, the virus particles, purified on a sucrose gradient, are pelleted by ultracentrifugation and resuspended in a lysis buffer containing 1% Triton X-100 and 1% lysolecithin. The released RNPs are then purified by glycerol gradient centrifugation. Figure II.3A shows an analysis of gradient fractions by denaturing gel electrophoresis (SDS-PAGE). The RNP peak fractions, characterized by NP and polymerase bands on the gel (figure I.3A, lanes 7-9), are pooled and concentrated before modification experiments. The integrity of the RNA in RNP preparations can be analyzed by electrophoresis on 6% polyacrylamide gels (figure I.3B). The genomic RNA segments of different lengths have been visualized in this case by ethidium bromide staining. Only 6 bands are readily visible on the RNA gels, because RNA segments 1-3 migrate together in the topmost band due to their very similar sizes. *In vitro* transcribed segment 8 RNA (890 nt) served as a size standard on this gel.

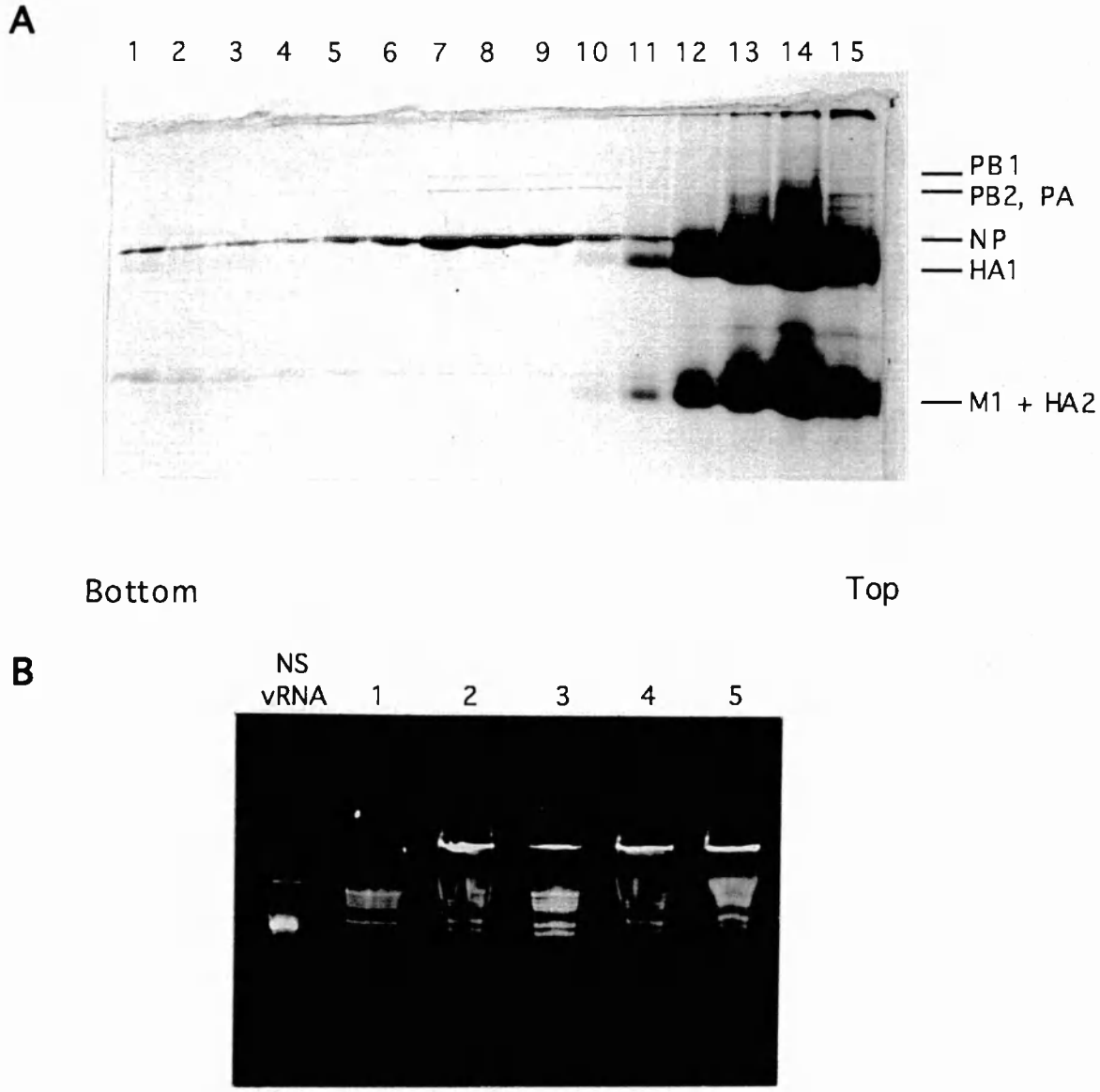


Figure I.3: Glycerol gradient purification of influenza virus RNPs from detergent disrupted virus. Viral proteins in fractions collected from the bottom of a 30-60% continous glycerol gradient were analyzed by 12% SDS-PAGE and Coomassie staining (A). RNPs in fractions 7-9 were pooled and analyzed for RNA content by 6% urea-PAGE and EtBr-staining (B, lane 1). Lanes 2-5 show vRNA samples from different RNP preparations. NS vRNA shows in vitro transcribed segment 8 vRNA

I.3.2. Analysis of DMS modification patterns on RNPs

The accessibility of bases near the 5' end of the RNP structure was studied with the segment 8 RNP coding for the NS proteins. The vRNA on the RNPs is readily modified by DMS as can be seen on figure I.4, which shows a set of DMS modification reactions comprising control (C) and three reactions with increasing amounts of DMS.

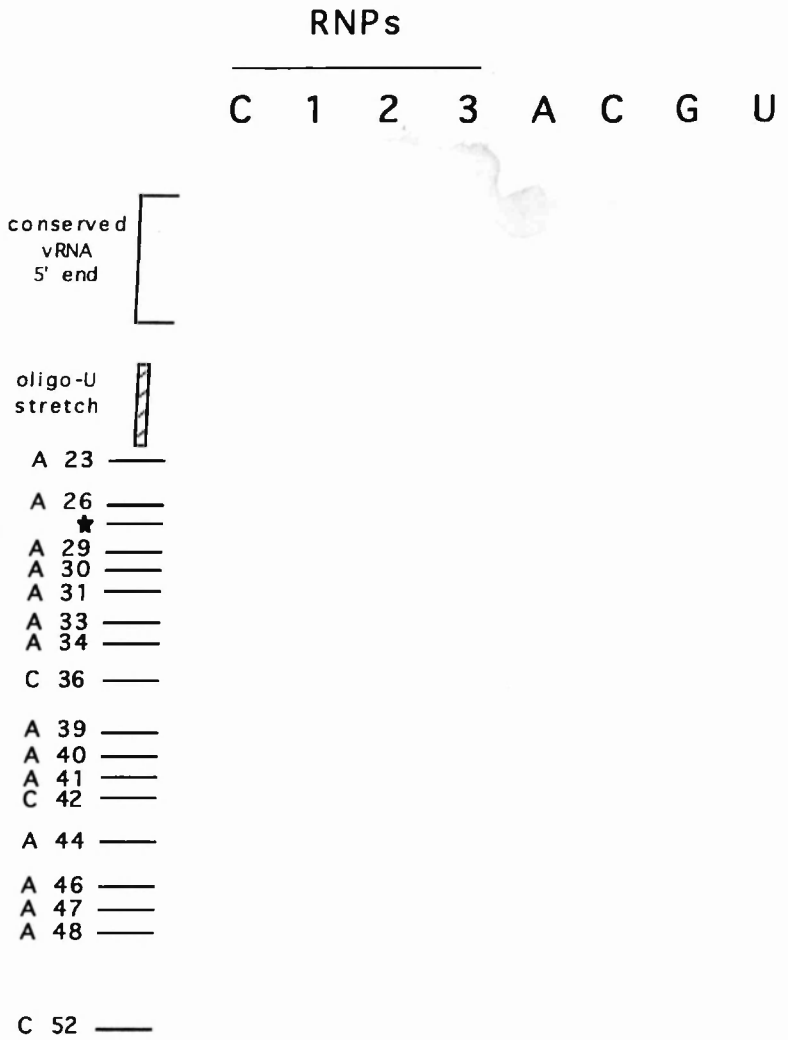


Figure I.4: 16% PAGE autoradiogram of the cDNA fragments produced after reverse transcription of DMS modified RNPs using a primer specific for segment 8 vRNA. The unreactive bases at the 5' end of the vRNA are indicated by a bracket. The reactivity of all As and Cs of the vRNA are shown from nucleotides 23 to 52. The star indicates nucleotide 27, where the sequence of our virus preparation deviates from the published NS vRNA sequence of influenza A/PR/8/34. Lane C is an incubation control of unmodified RNPs. Lanes 1,2 and 3 represent incubation of RNPs with 0.1, 0.2 and 0.6μl DMS respectively. Lanes ACGU are vRNA dideoxy-sequencing reactions of segment 8 vRNA.

On the right side of the gel is a RNA sequencing reaction (ACGU) to identify the modified positions on the vRNA sequence. The modified positions have been resolved by the reverse transcription technique (see I.2). All A and C residues downstream from A23 are reactive at their Watson-Crick positions and appear as bands on the analytical gel. This indicates that NP binding to vRNA does not involve the Watson-Crick positions of bases. The RNA bases are freely accessible from the solvent. The residues near the top of the gel

corresponding to the conserved 5' end of the vRNA, in particular A4, A6, A7, A8, C9, A10 and A11, indicated by the square bracket, were not modified. This base protection could in principle partially be caused by base-pairing of the conserved ends of the vRNA as has been observed on small model RNAs. Figure I.5 shows the secondary structure, that is formed between the conserved influenza virus vRNA ends on a naked, 84 nucleotide model RNA in solution according to Baudin et al. 1994.

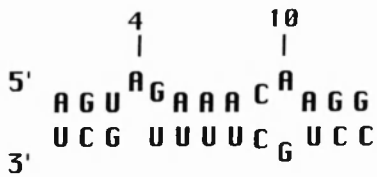


Figure I.5: Partially base-paired secondary structure ("panhandle") formed between the ends of segment 8 vRNA on small, naked model RNA in solution. Adapted from Baudin et al. (1994).

Of particular interest are the adenines A4 and A10, that are located in bulges and have been shown to be exposed in this RNA secondary structure. The fact that they prove completely unreactive on native RNPs strongly suggests an interaction of the polymerase protein with the 5' end of the vRNA. To directly demonstrate this interaction and to possibly identify polymerase subunits responsible for vRNA binding, we searched for means to selectively remove polymerase subunits from the RNPs.

I.3.3. Removal of the polymerase from RNPs

Polymerase binding to RNPs had been found to be sensitive to incubation with deoxycholate (DOC) (Inglis et al., 1976). The effect of increasing amounts of DOC on the composition of RNPs was analyzed by SDS-PAGE and it was confirmed, that DOC was capable of preferentially dissociating polymerase proteins from RNP structures. No changes in the sedimentation pattern of the RNPs were observed on glycerol gradients, that were used to separate the residual RNPs from detached polymerase proteins. Figure I.6 shows two silver-stained gels of a SDS-PAGE analysis after a succession of glycerol gradients to prepare polymerase-free RNPs. The top gel shows a typical separation of viral envelope and M1 proteins from complete RNPs on 30 ml continuous glycerol gradients similar to that shown in figure I.3A. Polymerase-free RNPs were obtained after incubation in 1 % DOC and recentrifugation on either a 12 ml glycerol step gradient (figure I.6B), in order to concentrate the RNPs on top of a 70% glycerol cushion, or a 30-60% glycerol

continous gradient, where the polymerase-free RNPs sedimented to regions of roughly 40% glycerol similar to complete RNPs (not shown).

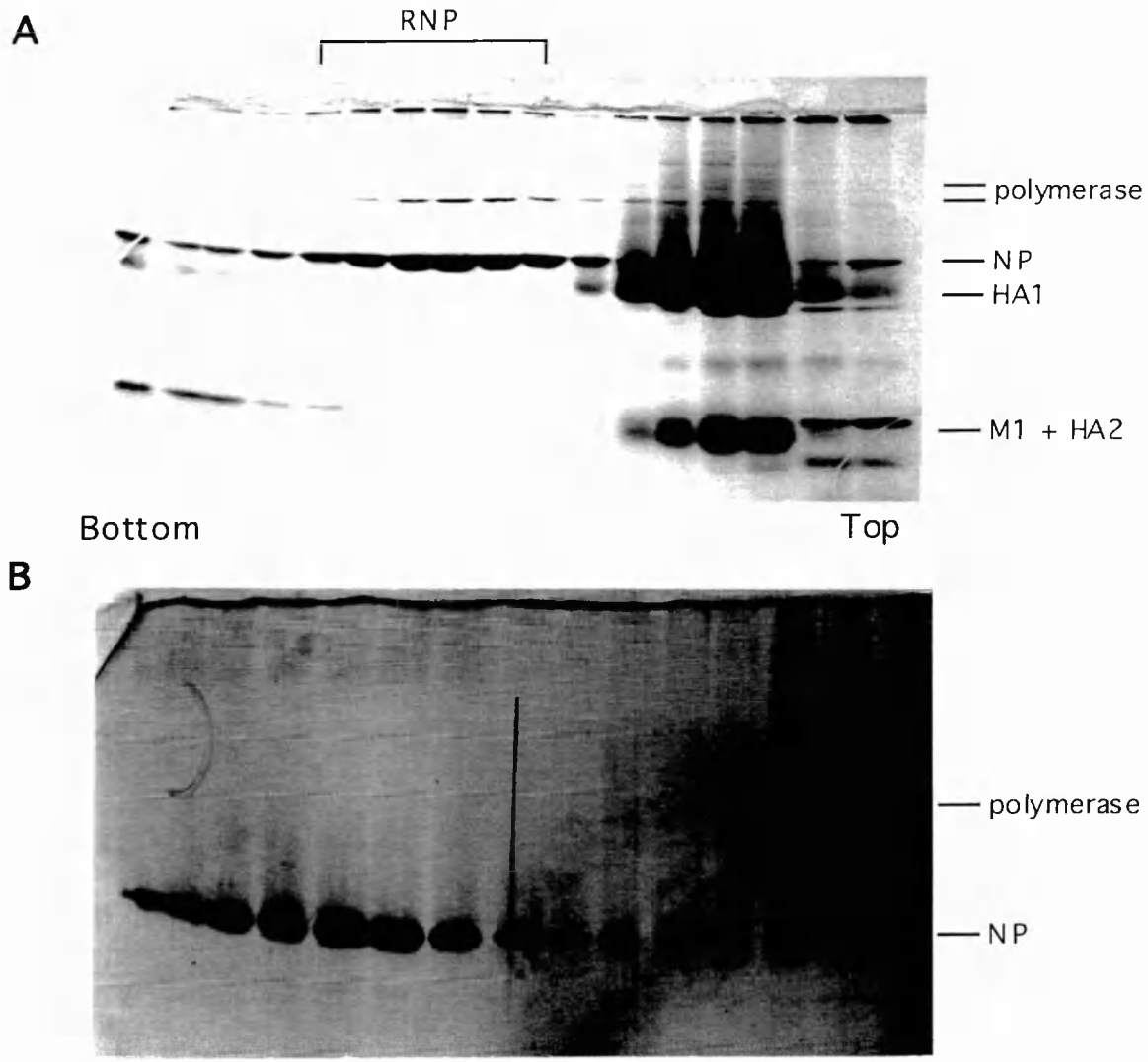


Figure I.6: (A), 12% SDS-PAGE of fractions from 30 ml 30-60% glycerol gradients for the isolation of complete RNPs. Fractions indicated by the bracket were pooled, concentrated, treated with 1% DOC and loaded onto a 70%, 30%, 20% glycerol step gradient, the fractions of which were analyzed in the same way. Both gels were silver stained.

The efficiency of polymerase removal was dependent on the concentration of RNPs. At 1 mg/ml RNP, even elevated concentrations of DOC did never completely remove the polymerase from the RNPs. The DOC treated RNPs were centrifuged through a 25 % glycerol cushion, and the pellet and supernatant fractions were analyzed by SDS-PAGE (figure I.7). A fast migrating polymerase subunit was dissociating from RNPs at lower DOC concentrations than the others. This subunit was identified as PA by Western blot analysis, but it could not be excluded, that both PA and PB2 dissociated simultaneously under these conditions. Complete and selective removal of polymerase proteins as shown

in figure I.6 was best achieved with 0.1 mg/ml of RNPs in 1 % DOC. Intermediate concentrations of DOC (0-1 %) did reproducibly and preferentially remove the faster migrating polymerase band from the RNPs. However, according to Western blot analysis, complete, selective removal of only the PA protein could not be achieved. PA started to dissociate at 0.025 % DOC, but was only below detection level on the RNPs after 1 % DOC incubation, when the other subunits had also dissociated. We did not succeed in the preparation of PA-free RNPs with this method. The efficacy of DOC was presumably related to the ionic nature of the detergent, because the RNPs remained complete even in the presence of up to 2 % CHAPS.

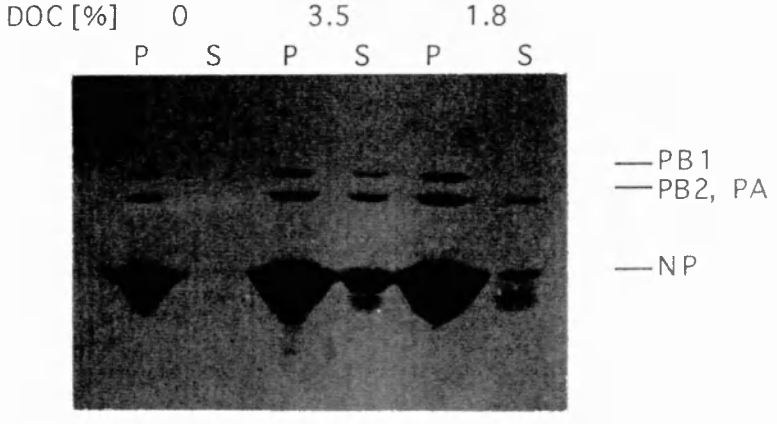


Figure I.7: 12% SDS-PAGE of RNPs incubated at 1 mg/ml with different concentrations of DOC, as indicated above the gel, then centrifuged through a 25% glycerol cushion. One of the smaller polymerase subunits is separated from RNPs at lower DOC concentrations than the others. DOC dependent dissociation of polymerase subunits from RNPs is dependent on RNP concentration. "P" and "S" stand for pellet and supernatant fractions.

I.3.4. Comparison of modification patterns: RNP versus RNP -pol

Figure I.8 shows a DMS modification experiment with RNP, similar to that described above (figure I.4), compared with the modification of RNPs devoid of polymerase (RNP -pol). No changes are observed in the NP binding pattern. All A and C residues downstream of position 23 are reactive with DMS in both cases. However, the residues at the 5' end of the vRNA, that are protected on complete RNPs become reactive after removal of the polymerase complex.

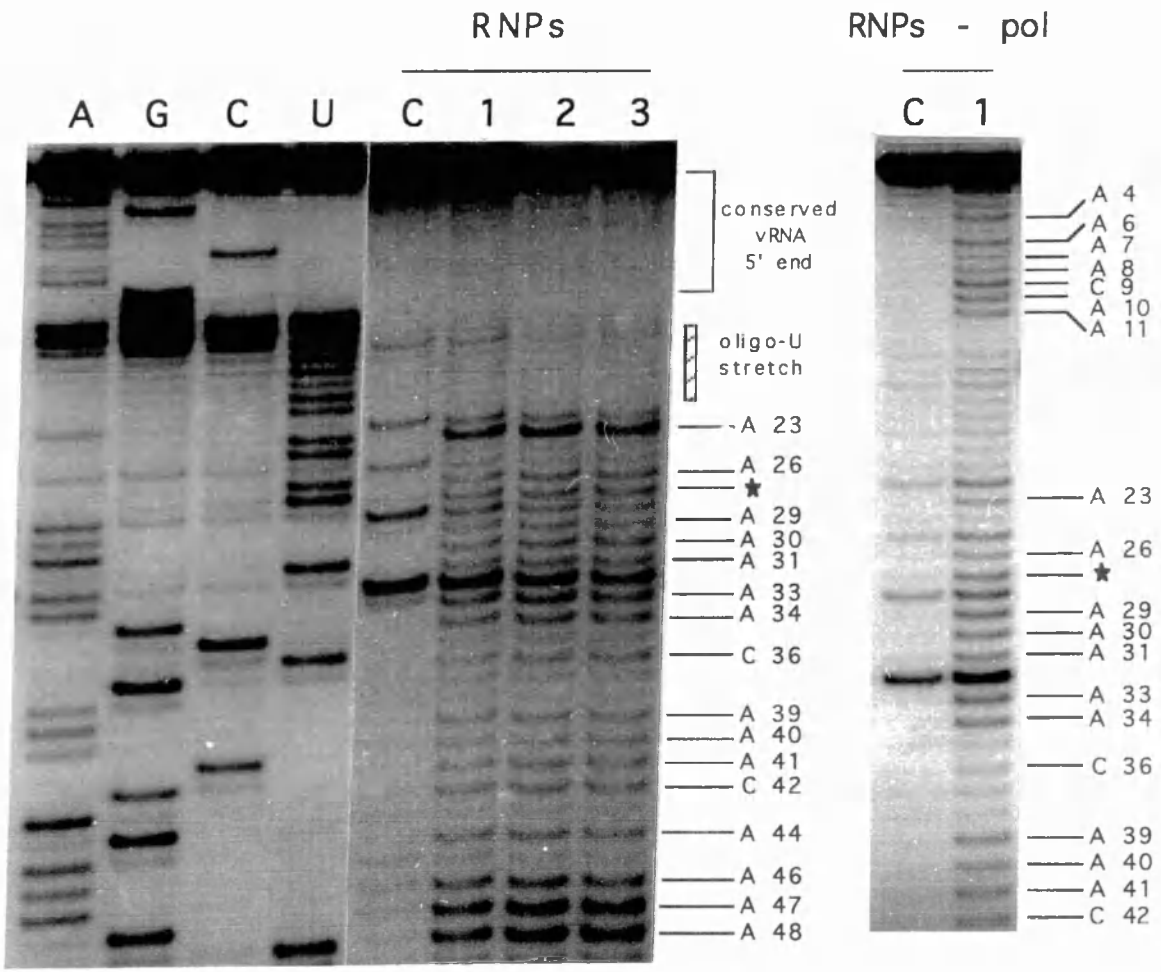


Figure I.8: 16% PAGE analysis of the cDNA products from DMS modified RNPs and RNPs lacking the polymerase complex (RNPs-pol). The reactive bases are indicated on the right. Lane C is an incubation control of unmodified RNPs. Lanes 1-3 are incubations of RNPs with 0.1, 0.2 and 0.6 μl DMS, lanes AGCU are RNA dideoxy-sequencing reactions.

Similar information, but now on the guanosines, was obtained by the analysis of RNPs by kethoxal modification, specific for N1 and N2 of G (figure I.9) and RNAse T1 digestion, specific for unpaired G (figure I.10). On complete RNPs, the guanines G5, G12, G13 and G14 were not accessible for either kethoxal or RNAse T1. The protection of G5 was again a direct evidence for the presence of the polymerase complex at the 5' end,

because this base shows significant reactivity on the model panhandle structure of naked RNA (figure I.11), which suggests that the observed G5 protection was not due to a RNA secondary structure. The reactivity of base G16 in the segment 8 sequence could not be determined because of modification unrelated stops of the reverse transcriptase in this region. The next G residue in the segment 8 vRNA sequence is located at position 35. G35 and all following G residues were reactive to Watson-Crick specific modification on RNPs.

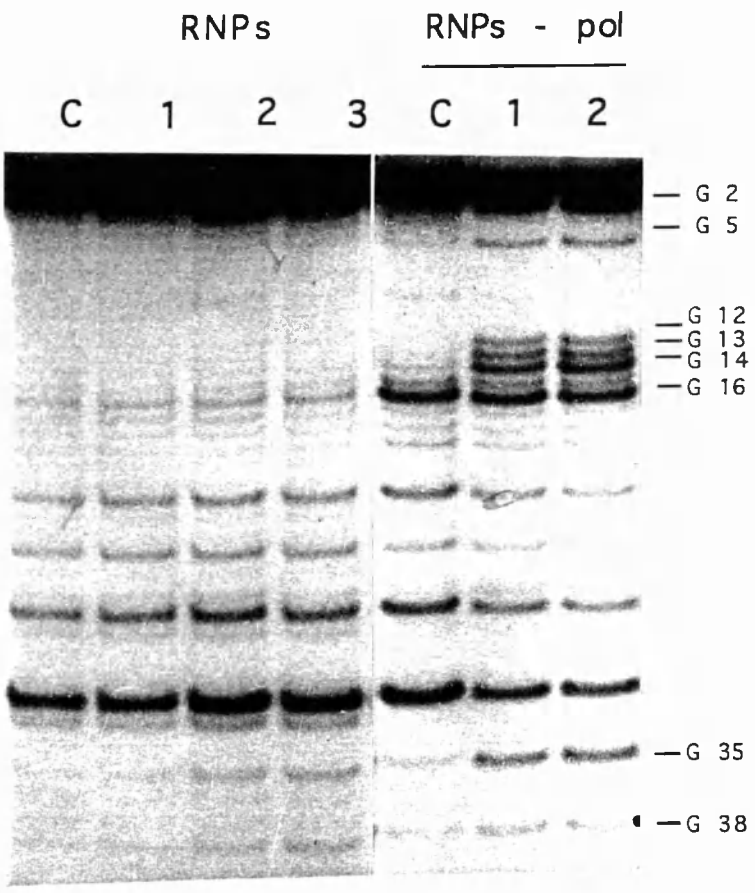


Figure I.9: 16% PAGE analysis of the cDNA products from kethoxal modified RNPs and RNPs lacking the polymerase complex (RNPs-pol). The reactive bases are indicated on the right. Lane C is an incubation control of unmodified RNPs. Lanes 1-3 are incubations of RNPs with kethoxal for 5, 10 and 20 min. RNPs-pol were incubated 20 and 60 min, lanes 1 and 2.

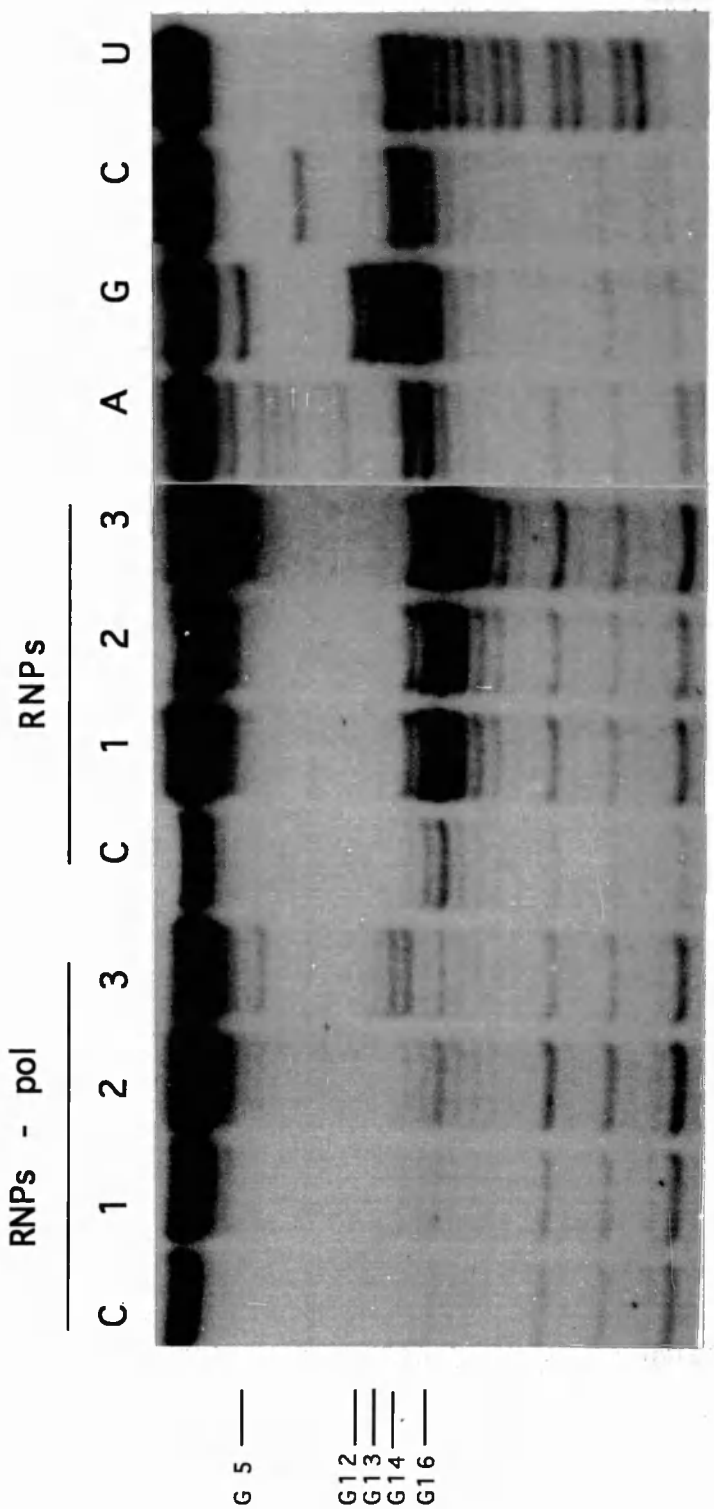


Figure I .10: 16% PAGE autoradiogram of the segment 8 derived cDNA products produced after reverse transcription of T1 RNAse digested RNPs and RNPs lacking the polymerase complex (RNPs-pol). The reactive bases are indicated on the right. Lane C is an incubation control of unmodified RNPs. Lanes 1-3 are incubations of RNPs with 0.1, 0.5 and 1 U of RNAse T1 respectively. Lanes AGCU are vRNA dideoxy-sequencing reactions performed with AMV-RT.

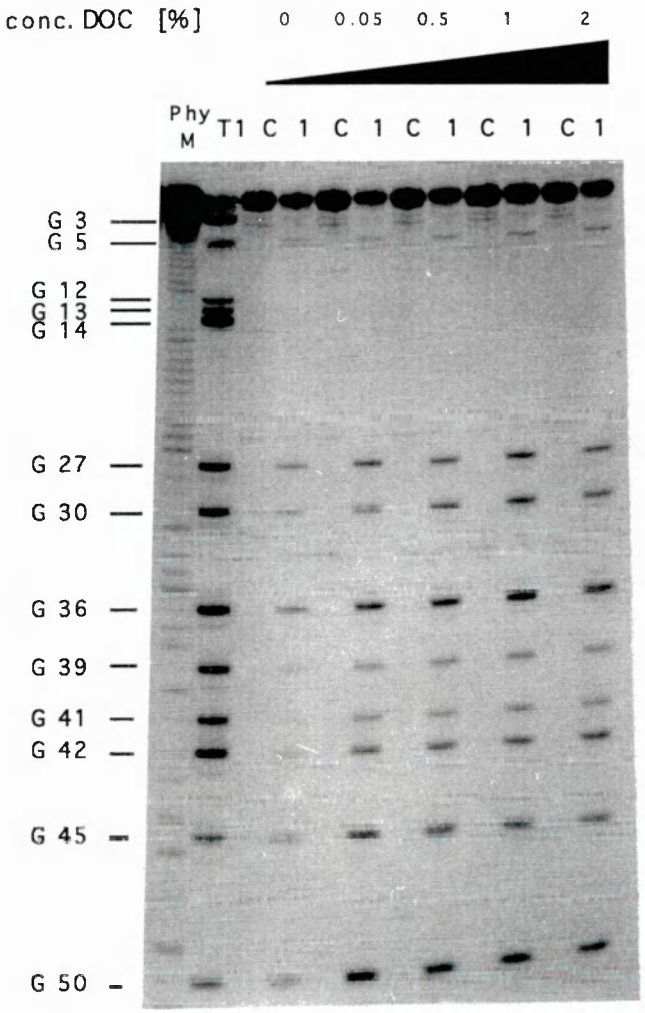


Figure I.11: RNAse T1 digestion of a 3' end labeled, 81 nt panhandle model RNA in the presence of increasing concentrations of deoxycholate (DOC). The positions of G residues are indicated on the left. G12, G13 and G14 are protected from hydrolysis by base-pairing with C residues at the 3' end of the molecule. The base-pairing interactions are not disturbed in the presence of DOC. PhyM and T1 denote sequencing reactions of the panhandle RNA with the corresponding RNAses under denaturing conditions performed with 1 U and 0.5 U of enzyme respectively. Lanes "C" are incubation controls without RNAse. Lanes "1" are RNAse T1 digestions of panhandle RNA under native conditions with 0.5 U of enzyme.

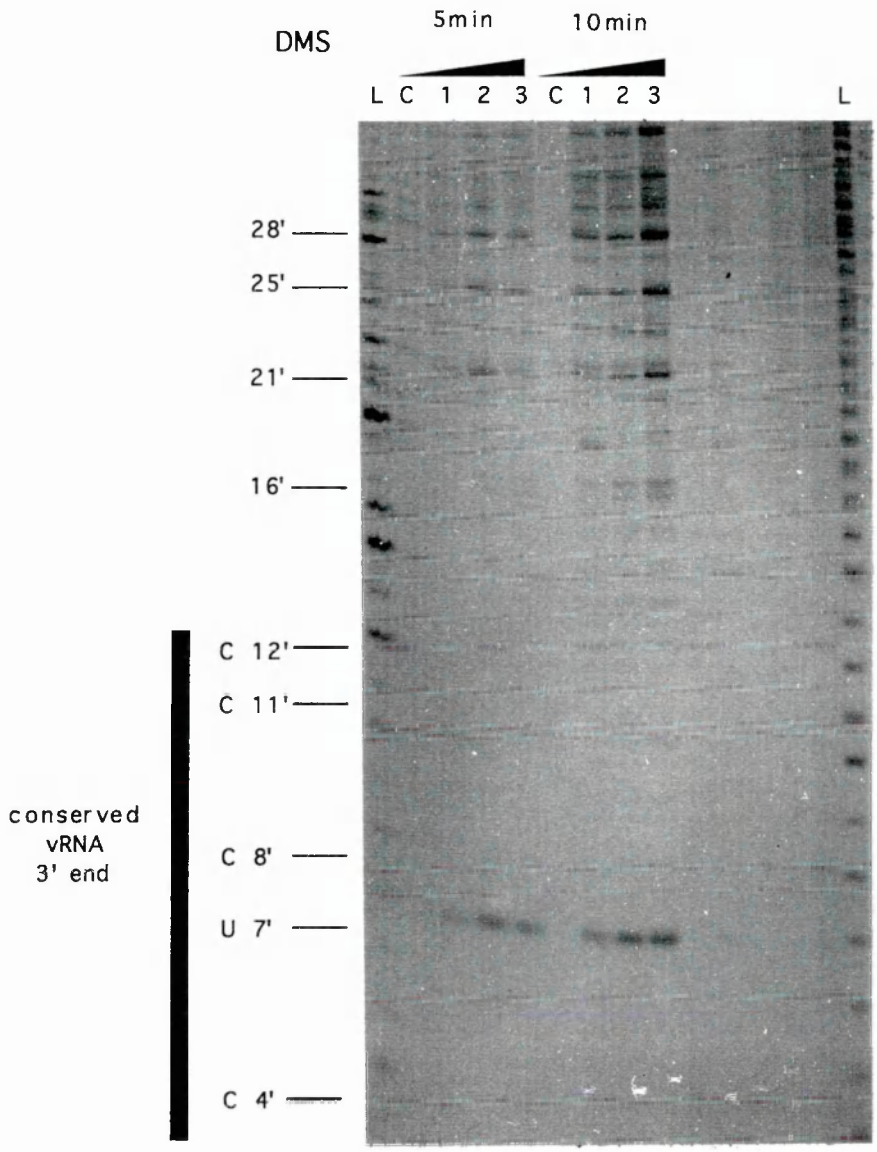


Figure I.12: Autoradiograph of chemical cleavage reactions of a mixture of all 3' end labeled vRNAs after DMS modification of complete RNPs. The position of the conserved 3' end on the sequence is shown by the black bar and all C residues in this region are indicated on the left (C4' to C12'). Further upstream, the sequence of the different genome segments diverts, and strongly reactive positions have been marked by numbers according to their position in the sequence. The autoradiograph depicts two sets of DMS reactions performed on RNPs for 5 min and 10 min respectively, each time with increasing amounts of DMS. Lanes "L" are alkaline hydrolysis ladders of vRNA, lanes "C" incubation controls without DMS. Lanes 1-3 are incubations of RNPs with 0.2, 0.5 and 1 μl DMS respectively.

I.3.6. Compilation of base reactivities on Influenza virus RNPs

Figure I.13 gives an overview of the modification data from complete RNPs (top) and polymerase-free RNPs (bottom) compared to data obtained with naked, panhandle model RNA. Circles and squares around the bases indicate the reactivity of Watson-Crick positions of the bases on RNPs, stars and cubes above and below the bases the reactivity of bases on perfectly base-paired, small, naked RNA as indicated in the figure. The absence of circles or squares means that the reactivity could not be determined. The pattern of base modification on RNPs is different from naked panhandle RNA and from in vitro reconstituted NP-RNA complexes (Baudin et al., 1994) at the vRNA ends. The observation, on complete RNPs, of protected bases, that are accessible for modification on optimally base-paired panhandle structures, together with the fact that NP binding to vRNA does not involve the Watson-Crick positions of bases, indicates that the polymerase protein is bound to the conserved 5' end of the vRNA on native RNPs. The 3' end is part of a ternary complex together with the polymerase and the 5' end, resulting in the protection of bases until position 15' on the 3' end and at least position 14 on the 5' end. Downstream of the polymerase binding site the bases are exposed and highly reactive to all modification reagents tested. In the absence of the polymerase the vRNA ends do not interact with each other and the RNA is devoid of any secondary structure. The extent of base protection from modification, which is observed in the presence of the polymerase, correlates well with the boundary of the theoretical panhandle structure, that can in principle be formed by base-pairing of the vRNA ends. Depending on virus strain and RNA segment, 12-17 nucleotides from the 3' end could in theory be annealed with the corresponding 5' ends (see figure I.14 and (Robertson, 1979; Skehel and Hay, 1978; Stoeckle et al., 1987). This suggests that the polymerase interacts with both conserved vRNA ends and this process may lead to the annealing of the complementary sequences.

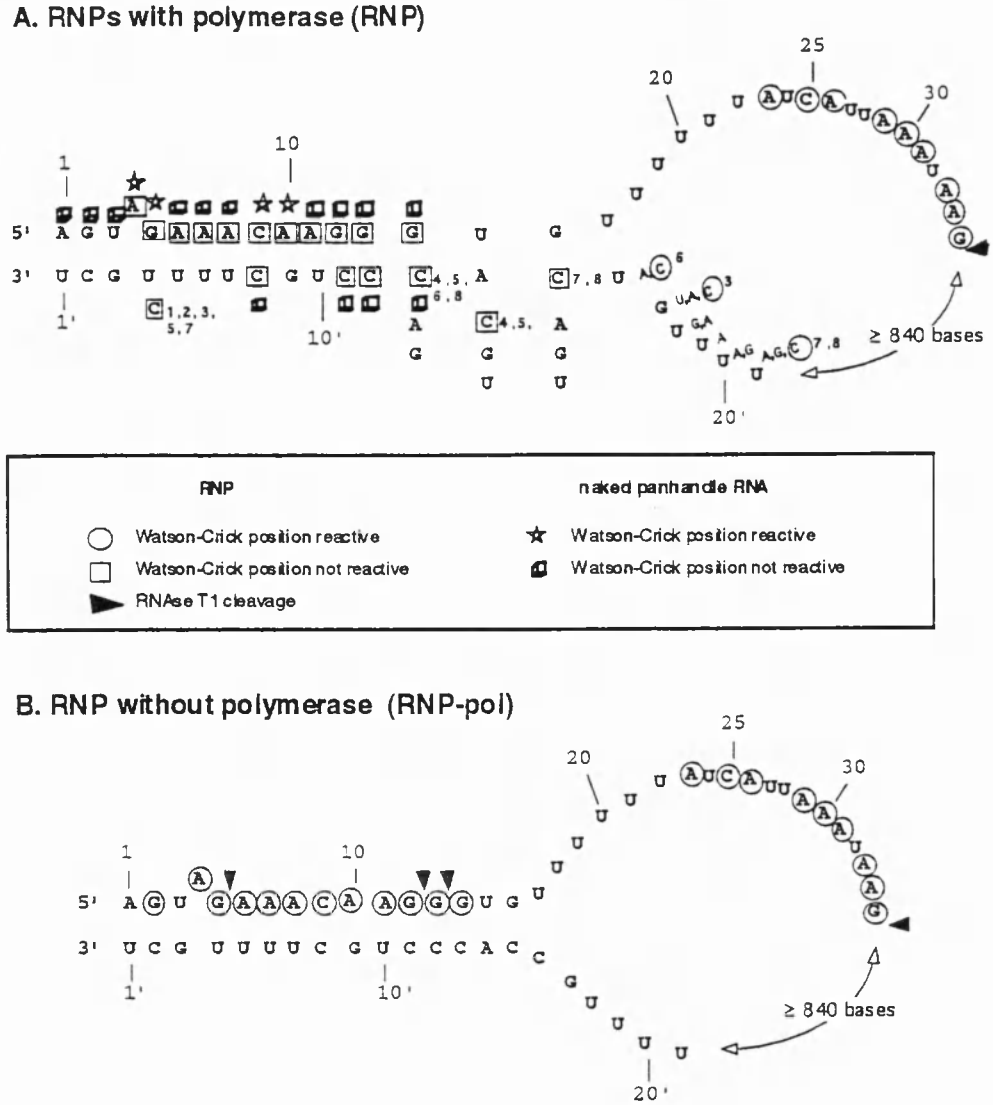


Figure I.13: Compilation of modification data from complete RNP (top) and polymerase-free RNP (bottom) compared to data obtained with a 84 nt, in vitro transcribed panhandle RNA. Circles and squares around bases indicate the reactivity of Watson-Crick positions of the bases on RNP. Stars and cubes above and below bases indicate reactivity of bases on naked, annealed panhandle RNA. The absence of circles or squares means that the reactivity could not be determined. The cleavage by RNAse T1 is indicated by arrowheads. The scheme is drawn with the sequence of segment 8 vRNA in the center. The base reactivities at the 3' end have been determined in a mixture of all segments. The sequence deviations in the different RNA segments of influenza A/PR/8/34 are shown between nucleotides 13'-21' of the 3' end. The numbers denote the segments, that carry the respective cytosine at the specific position of the sequence. The nucleotides 13'-15' are usually, but not always, complementary to nucleotides 14-16 of the 5' end (compare figure I.14). The sequences have been extracted from the GENBANK/EMBL database.

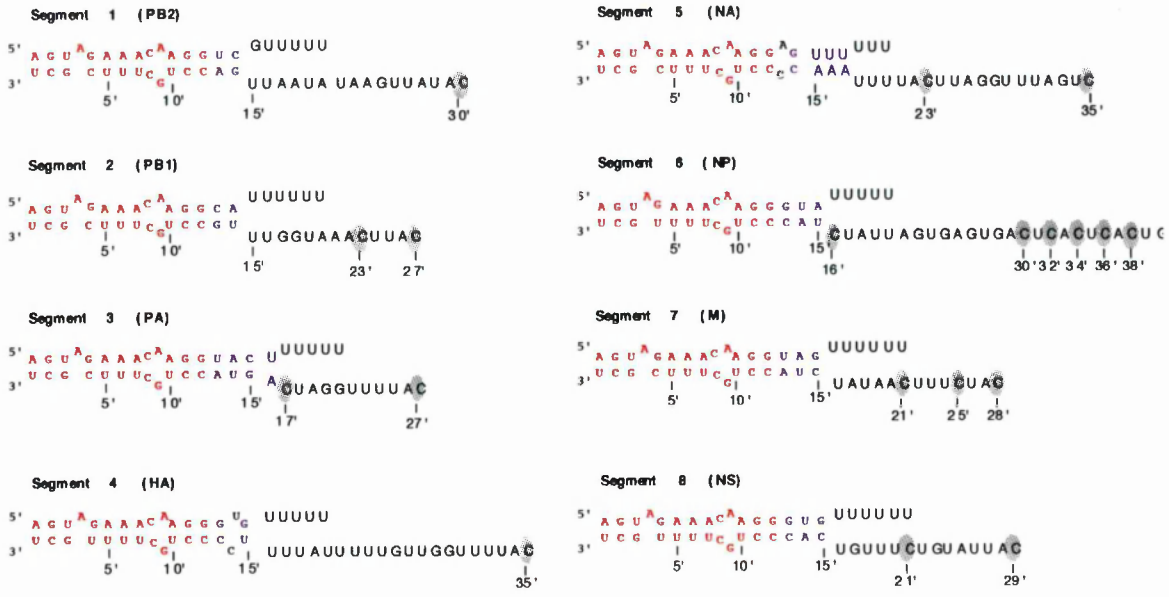


Figure I.14: Sequences at the ends of Influenza A/PR/8/34 virus show a partial, inverted complementarity. The conserved sequences are in red characters. Blue letters show the extension of complementarity into the non-conserved sequence. The sequences at the 5' end are shown until the oligoU transcription termination signal. The 3' end sequences are shown until the UAC, which encodes the translation start codon. C residues, that have been reactive on RNPs are accentuated by a grey background. The sequences have been extracted from the GENBANK/EMBL data base.

I.4 Discussion

The polymerase is the central protein complex of the influenza virus replication cycle. It is bound to the genomic RNPs in the virion, but the binding site has only been poorly defined (Honda et al., 1987; Murti et al., 1988). After the RNPs have entered the cell nucleus, transcription of viral mRNA starts from the 3' ends of the vRNA templates and terminates at an oligoU stretch near the 5' end of the vRNA. Later in infection the polymerase generates full-length complementary transcripts (cRNA), which serve as templates for the production of secondary, genomic vRNAs. There is still considerable uncertainty concerning the nature of the cis-acting sites and the mechanisms, that are involved in the regulation of the replicative processes. All necessary signals for replication and genome packaging seem to reside in the terminal sequences of the vRNA, which are partially complementary (Luytjes et al., 1989), and several lines of evidence have implied a crucial regulatory role for a double stranded panhandle structure in viral transcription (Cianci et al., 1995; Fodor et al., 1995; Hsu et al., 1987; Li and Palese, 1994).

I.4.1 Binding of nucleoprotein and polymerase to the genomic RNA

The results outlined above suggest that on the RNP the polymerase forms a ternary complex including both conserved ends of the viral RNA. Direct interactions of the polymerase with Watson-Crick positions at the 5' end could be inferred from RNA modification analysis and the base protection observed at the 3' end indicates that it is part of the complex. However, the studies on complete RNPs alone could not clearly determine if the vRNA 3' end was held in the complex mainly by the polymerase protein or by base-pairing interactions with the 5' end. This problem has been approached by studying RNPs from which the polymerase was selectively separated. Both vRNA ends become completely exposed and single-stranded after removal of the polymerase by the detergent DOC. Furthermore, DOC itself has no effect on RNA secondary structures in solution and it is therefore unlikely, that DOC would interfere with a possible panhandle structure formed by the vRNA ends in RNPs. The separation of the vRNA ends is most likely a direct consequence of the absence of the polymerase. Even at very low DOC concentrations (0.1-1%), that only partially remove the faster migrating polymerase band from the RNPs, the vRNA ends in RNPs already became reactive. Additionally, very low DOC concentrations already strongly inhibited *in vitro* transcription activity, before any dissociation of protein could be observed on SDS-PAGE gels (figure I.14-2, page 33). Therefore, DOC does not interfere with RNA secondary structure, but with protein interactions in the polymerase complex and between polymerase and vRNA.

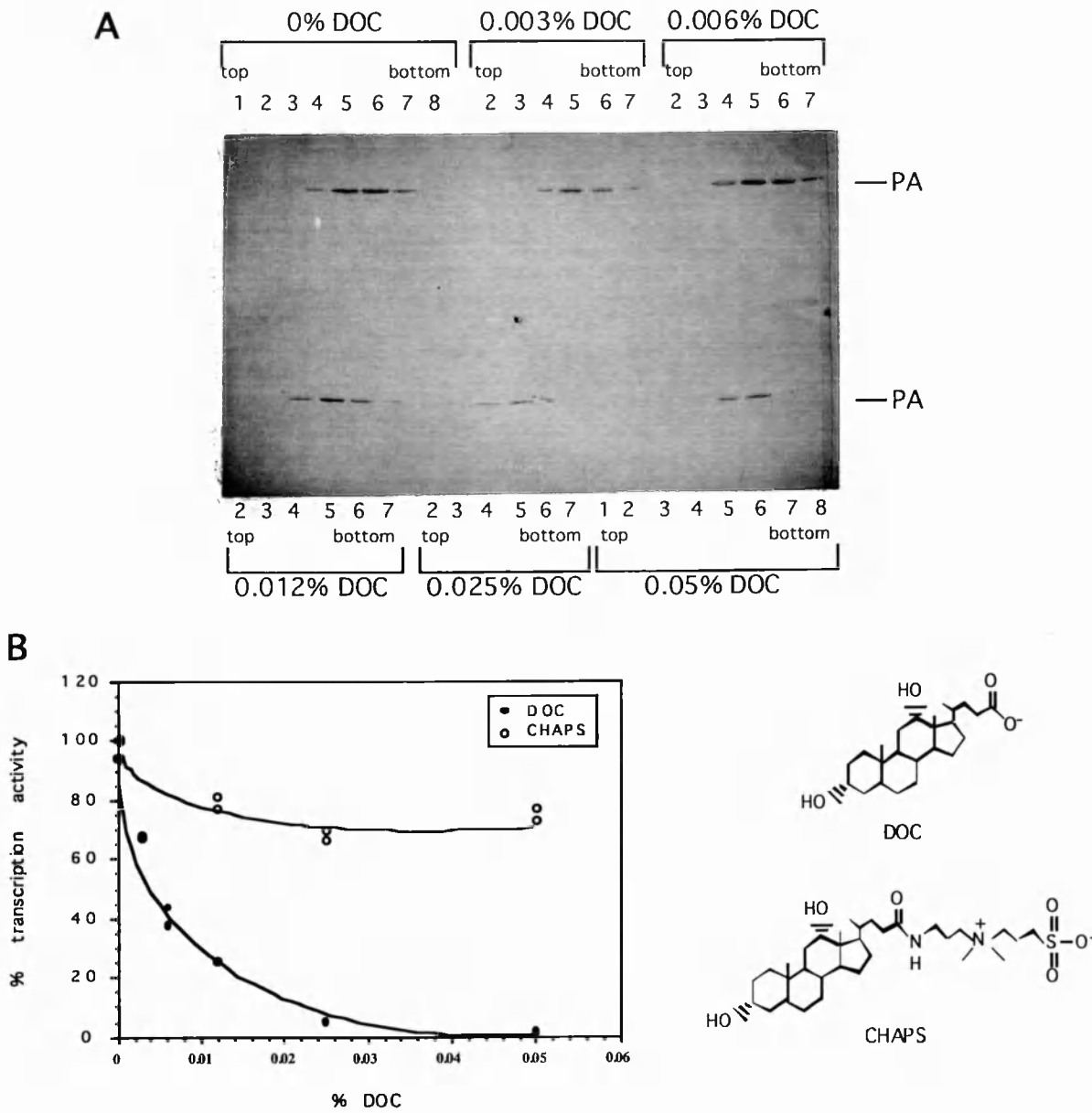


Figure I.14-2: Effect of low concentrations of sodium deoxycholate (DOC) on the influenza virus polymerase complex. RNPs have been incubated with the indicated concentrations of DOC and either loaded onto a 500 µl glycerol step gradient (25, 30, 40, 50, 60 and 70%) and analyzed by Western blot (A), or employed in ApG primed transcription assays (B). In (A), six 12% SDS-PAGE gels have been blotted onto the same nitrocellulose membrane and developed together with L35E12 monoclonal anti-PA antibody. In (B), transcription activity was determined by counting the amount of nucleotides incorporated into TCA-insoluble material. Incorporation of 400 pmol UTP label per pmol RNP in 1 hour was set as 100% activity. Transcription activity of RNP is already maximally inhibited by DOC, when the majority of the polymerase is still bound to the RNP.

The spontanous formation of a panhandle structure has been observed on small (<100 nt) model RNAs in solution (Baudin et al., 1994) and reproduced in this study (figure I.11). The genomic vRNAs, on the other hand, are much longer, i.e. between 890 and 2341 nucleotides length. Modification analysis of naked, in vitro transcribed segment 8 RNA (890 nt) showed that the vRNA ends were separated on these molecules and did not form a stable panhandle structure (figure I.15, courtesy of Dr. Florence Baudin). The bases in the region of the conserved ends were reactive with modifying agents and appear as stops of the reverse transcription reaction on the gel.

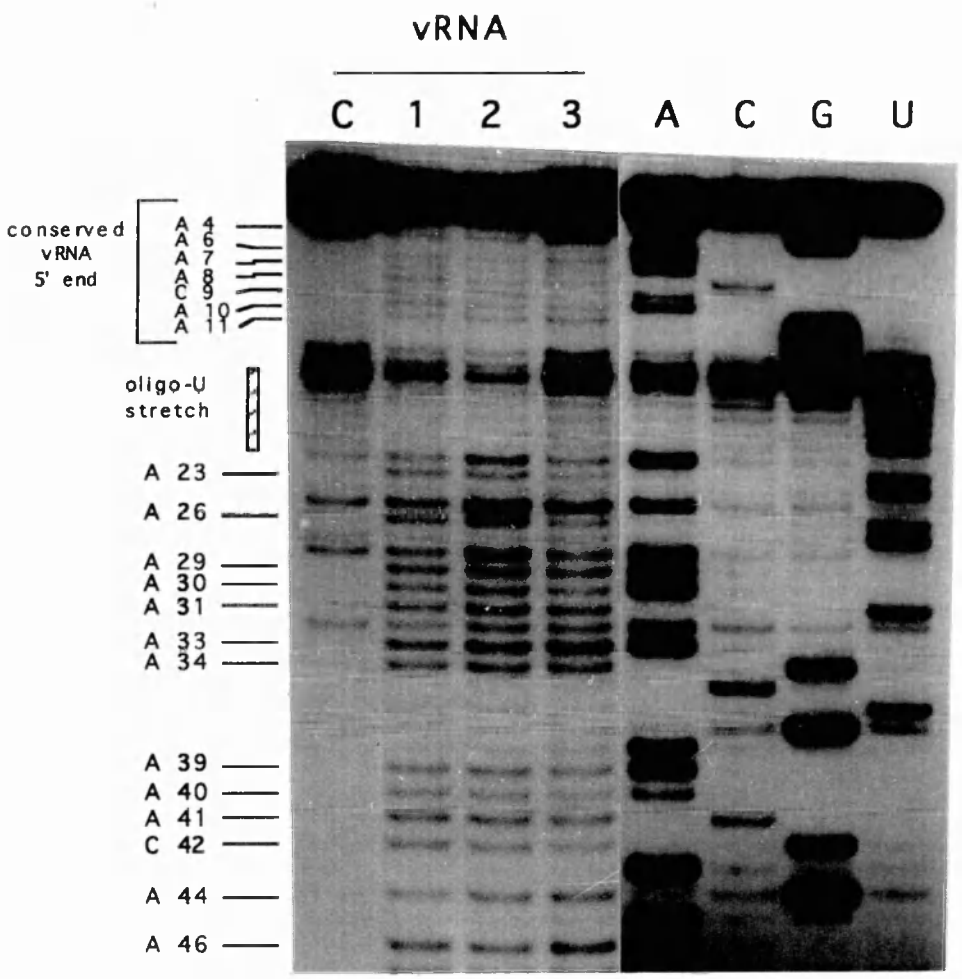


Figure I.15: 12% PAGE auoradiogram after DMS modification of naked segment 8 vRNA. The reactive bases are marked on the left. The position of the conserved 5' end on the sequence is indicated by a bracket, the position of the oligoU stretch by a hatched bar. Lane "C" is an incubation control of unmodified RNA. Lanes 1-3 represent reactions of RNA with 0.1, 0.2 and 0.6 μl DMS respectiely. Lanes ACGU are vRNA dideoxy-sequencing reactions.

This suggests that protein interactions are required to result in the extent of base-protection from modification as observed on complete RNPs. The absence of a panhandle structure on naked, genomic RNA has been observed before in psoralen crosslinking studies (Hsu et al., 1987). The RNA modification results clearly show that NP binding to the vRNA exposes the bases to the solvent. Downstream of the presumed polymerase protected site on the vRNPs the vRNA is held in a single-stranded conformation by NP. In addition, reconstitution experiments with purified NP and a small RNA containing the conserved influenza virus vRNA ends demonstrate that NP binding to RNA can result in the melting out of RNA secondary structures all along the RNA molecule (Baudin et al., 1994). NP can therefore be excluded as the protein that protects the vRNA ends from modification. On the other hand, the comparison of RNPs and polymerase free RNPs by modification analysis indicates that the polymerase holds the ends together. This clamping of the vRNA ends could be visualized by Dr. Rob Ruigrok using negative stain electron microscopy (figure I.16). The complete RNPs appear as relaxed, helical strands, that are wound back on themselves, often showing a loop at one or both ends (compare with Jennings et al., 1983). After removal of the polymerase complex, the strands come apart at one of the ends of the RNP structure, indicated by arrow heads in figure I.16. Otherwise the morphology is not changed by the DOC treatment. The alteration in the interaction at the ends becomes even clearer after the incubation of RNPs in high or low salt conditions. Surprisingly, in both cases the complete RNPs unwind to circular structures (figure I.16, middle and bottom left). Polymerase-free RNPs unwind to elongated structures under low salt conditions, but form very tightly packed structures in high salt (figure I.16, middle and bottom right). Both observations illustrate the loss of a restrictive contact between the vRNA ends caused by removal of the polymerase from the RNPs.

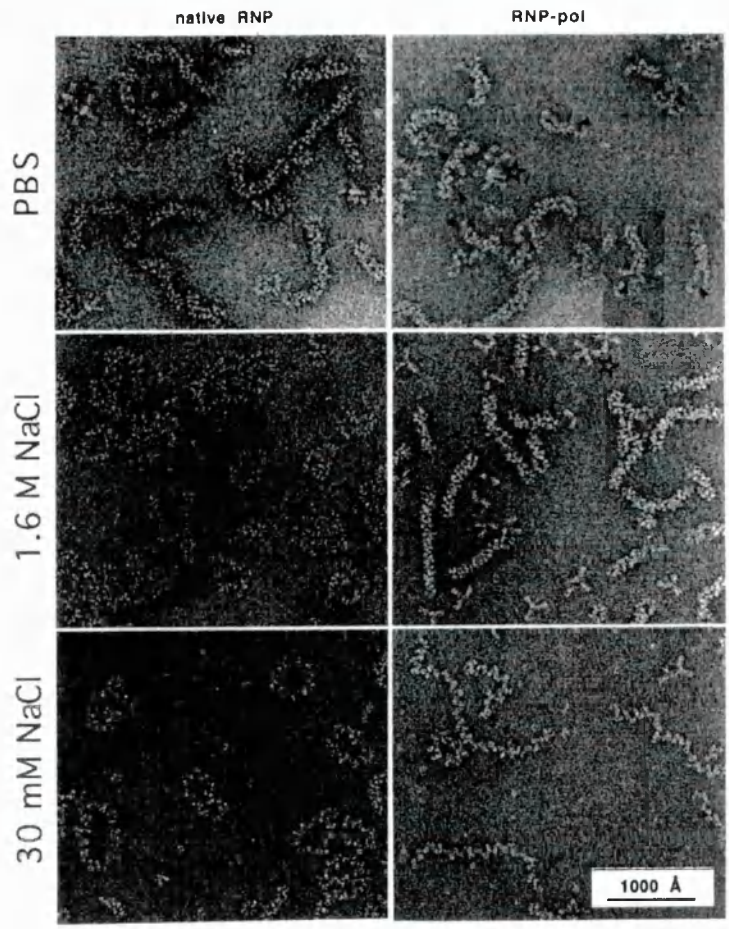


Figure I.16: Electron micrographs of negatively stained RNPs and RNPs-pol diluted and incubated with phosphate buffered saline (PBS), 5 times diluted in 2 M NaCl (1.6 M NaCl) or 5 times diluted in H_2O (30 mM NaCl). The stars indicate contaminating rosettes of hemagglutinin. The arrows in RNP-pol in PBS (top right) indicate where the ends of the RNPs have come apart.

The images of the conformational changes of salt-treated RNPs correlate well with their behaviour in glycerol gradients. Figure I.17A shows fractions from glycerol gradients of salt-treated RNPs. In 0.1 M salt, RNPs are found in the region of about 40-50% glycerol (top left, fractions 5 and 6). Under high salt conditions a large portion of the RNPs remain near the top of the gradient and do not pass through 30% glycerol. The RNPs that stay at the top exclusively retain the polymerase proteins and only these top fractions show specific transcription activity after dialysis, albeit at reduced level (figure I.17B). Therefore, they possibly represent the circular forms of the complete RNPs. Sometimes, the same RNP preparation contained both complete and polymerase-free RNPs. The latter could be recovered from fractions at the bottom of the gradient. They sedimented further into the gradient presumably due to the compact, polymerase-free high-salt structures. This means that according to virus preparations either a fraction of the RNPs can lose

their polymerase complex during RNP preparation, during incubation in high salt, or it might also be possible that sometimes polymerase-free RNPs are included into virus particles. Murti et al (1988) reported that about 30% of the RNPs in their preparations could not be labelled with anti-polymerase antibodies. We have also observed considerable variations in transcription activity depending more on virus batches than on RNP preparations (see chapter III). Using RNP preparations of low level transcription activity in RNA modification experiments we hardly observed complete protection at the vRNA ends and the background in the modification reactions was higher. The above shown RNA modification results were obtained with a virus preparation that showed almost maximal transcription activity in vitro (400 pmol UTP incorporated per pmol RNP per hour). All these observations together are consistent with the idea that the polymerase complex is responsible for holding the ends of the vRNA together by interacting with both termini on the viral RNPs. In the absence of the polymerase the ends are single-stranded and free to move and rotate after incubation in different salt conditions.

I.4.2 Implications of the RNP structure for transcription and replication

The remarkable conservation of the end sequences of the genomic RNA and their partial inverted complementarity suggested a possible function of a panhandle RNA secondary structure during virus transcription and/or replication, although partially complementary sequences at both ends are also expected, when the polymerase needs similar promoter sequences for plus and minus strand synthesis during replication (Skehel and Hay, 1978). The present results, suggesting the vRNA ends to be apart in the absence of the polymerase, seem to exclude the requirement of a preformed panhandle structure for polymerase binding to template RNA and rather support a model of sequential or independent binding to single vRNA ends, as has been proposed by Cianci et al. (1995). Considering the extent of base protection at the Watson-Crick positions that we observe on native RNPs, together with the fact that the vRNA ends on RNPs could be crosslinked by psoralen (Hsu et al., 1987), it is very likely that polymerase binding to the vRNA termini induces the formation of a certain amount of base-pairing. This agrees with in vitro transcription studies using mutant template RNA, which suggested that the formation of a so-called terminal RNA fork is a prerequisite for transcription initiation (Fodor et al., 1995). These authors propose that a few base pairs in the region of nucleotide 11 and downstream at the 5' end with the corresponding bases at the 3' end have to occur for transcription to be initiated in their system. They find no evidence for base-pairing with bases upstream of nucleotide 11 until the vRNA 5' end. This is consistent with the idea that the extreme 3' end of the vRNA has to be available in order to function as a single-stranded template during transcription initiation.

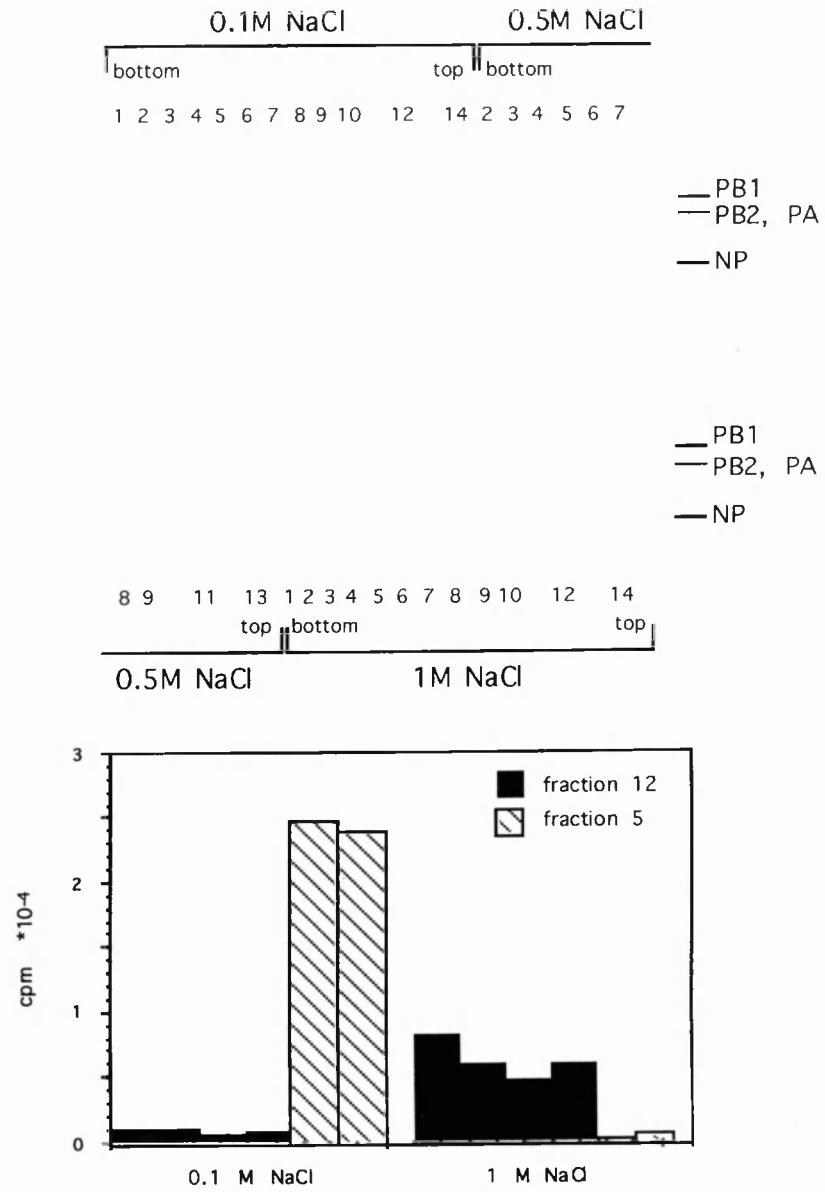


Figure I.17: Distribution of RNPs in glycerol gradients and influence of salt conditions. (A) shows two silver-stained 12% SDS-PAGE gels with fractions from three glycerol step gradients (1 ml 70% glycerol cushion, 500 µl steps of 60%, 50%, 40%, 30% glycerol in 0.1 M NaCl). RNPs were incubated 10 min at 37°C in different salt concentrations, as indicated, before being loaded onto the gradients, and centrifuged 2h at 240000 x g. Fractions of one time 1ml, then 250 µl, have been collected from the bottom. (B) depicts relative, ApG primed transcription activity of fraction 5 (from 2 experiments) and fraction 12 (from 4 experiments) counting the incorporation of radioactively labelled UTP into TCA insoluble material. In 0.1 M salt activity is detected in fraction 5, but not in fraction 12. After incubation in 1M salt, transcription activity can be recovered in fraction 12, but not in fraction 5. All fractions were dialysed against 0.1 M NaCl before the activity assay.

From the behaviour of the RNP proteins in RNA modification experiments, influenza virus RNP appears to be assembled from two antagonistic proteins: nucleoprotein activity favours the melting of RNA secondary structures and exposes the bases to the environment, whereas the polymerase complex anneals the two vRNA ends and

causes base protection. This antagonism constitutes an ideal arrangement for the regulation of a switch between a closed and an open RNP form, because in this situation such a switch only requires the manipulation of the fastening polymerase complex. An opening of the RNP is presumably needed for the production of full-length RNAs during replication, when the 5' end has to be freely accessible. On the other hand, during transcription, the 5' end of the vRNA template is not available since the mRNAs are incomplete transcripts terminated at the downstream oligoU stretch.

In electron microscopic studies the structure of complete RNP appears to be under a certain tension to be opened. Even under standard conditions, terminal loops of different sizes are often observed and the RNPs completely open up to circular structures after any changes of the salt concentration. This feature, that the RNP is poised to be unwound, resembles plasmid supercoiling in certain aspects and might be important for efficient transcription. The coiling of the influenza virus RNP has been found to be of opposite sense to the coiling of RNP from other negative strand RNA viruses (Heggeness et al., 1982). Moreover, the opening up of influenza virus RNP in high salt has been used in the past as an argument that influenza virus RNPs differ significantly from other negative strand RNPs, which form very tight structures in high salt and extended structures in low salt conditions (Heggeness et al., 1980). We find that in the absence of the polymerase, the influenza virus RNP behaves identical to the other RNPs in this respect. In the absence of constraint at the ends, the influenza virus RNPs extend in low salt and tighten up in high salt. To further follow this aspect of influenza virus RNP structure, that seems to be brought into tension by directed supercoiling and end clamping, it will be interesting to determine the direction of coiling on complete RNPs compared to polymerase-free RNPs and the factors involved in this mechanism.

Previously, a double-stranded RNA panhandle structure has been suggested to be important for transcription termination (Li and Palese, 1994; Luo et al., 1991). Several lines of evidence are in conflict with this interpretation. Our results suggest that, in the absence of the polymerase, the vRNA ends are not base-paired in RNPs. Even naked, full-length vRNA does not form a panhandle structure in solution according to RNA modification analysis (fig.4B), which is in agreement with the studies by Hsu et al. (1987), who could not crosslink the ends of protein-free vRNA with psoralen. These observations make it difficult to explain how the ends could stay base-paired with each other once the polymerase has left after transcription initiation. Nevertheless, all the structural studies on RNPs, all those on transcriptional mechanisms and even those on transcription termination are consistent with a model, where the formation of a partially base-paired RNA-fork, annealed by the polymerase complex, is necessary for transcription initiation only. This initial annealing of the ends by the polymerase prior to initiation would explain the dependence of transcription activity on a short, complementary sequence near the end. Because the theoretical panhandle RNA secondary structure appears to be unstable in the

absence of the polymerase, transcription termination at the oligoU stretch may rather be controlled by a regulatory protein binding to the conserved 5' end of the vRNA. This hypothetical, regulatory protein could participate in the switch between transcription and replication by determining the accessibility of the 5' end for being copied. Because RNA modification analysis shows that the polymerase interacts with the 5' end on native, viral RNPs, the polymerase itself or one of its subunits are prime candidates to harbour this regulatory function. In one possible model of the influenza virus replication cycle, extended from Cianci et al. (1995) and illustrated in figure I.18, the polymerase sequentially interacts with the 5' end and the 3' end of the viral RNA to form a functional transcription initation complex. The polymerase then advances along the vRNA template from the 3' end, but at the same time might stay bound to the 5' end. Transcription termination and polymerase stuttering at the oligoU stretch would then be imposed by sterical reasons. Replication is performed by a different form of the polymerase, that is only present in infected cells and has different RNA binding patterns (Beaton and Krug, 1986; Hay et al., 1980; Shapiro and Krug, 1988). This alternative form of the polymerase has no endonuclease activity, starts RNA synthesis without a primer RNA, releases the template 5' end and produces full-length replicates, which are concomitantly packaged by NP into RNP structures (figure I.18).

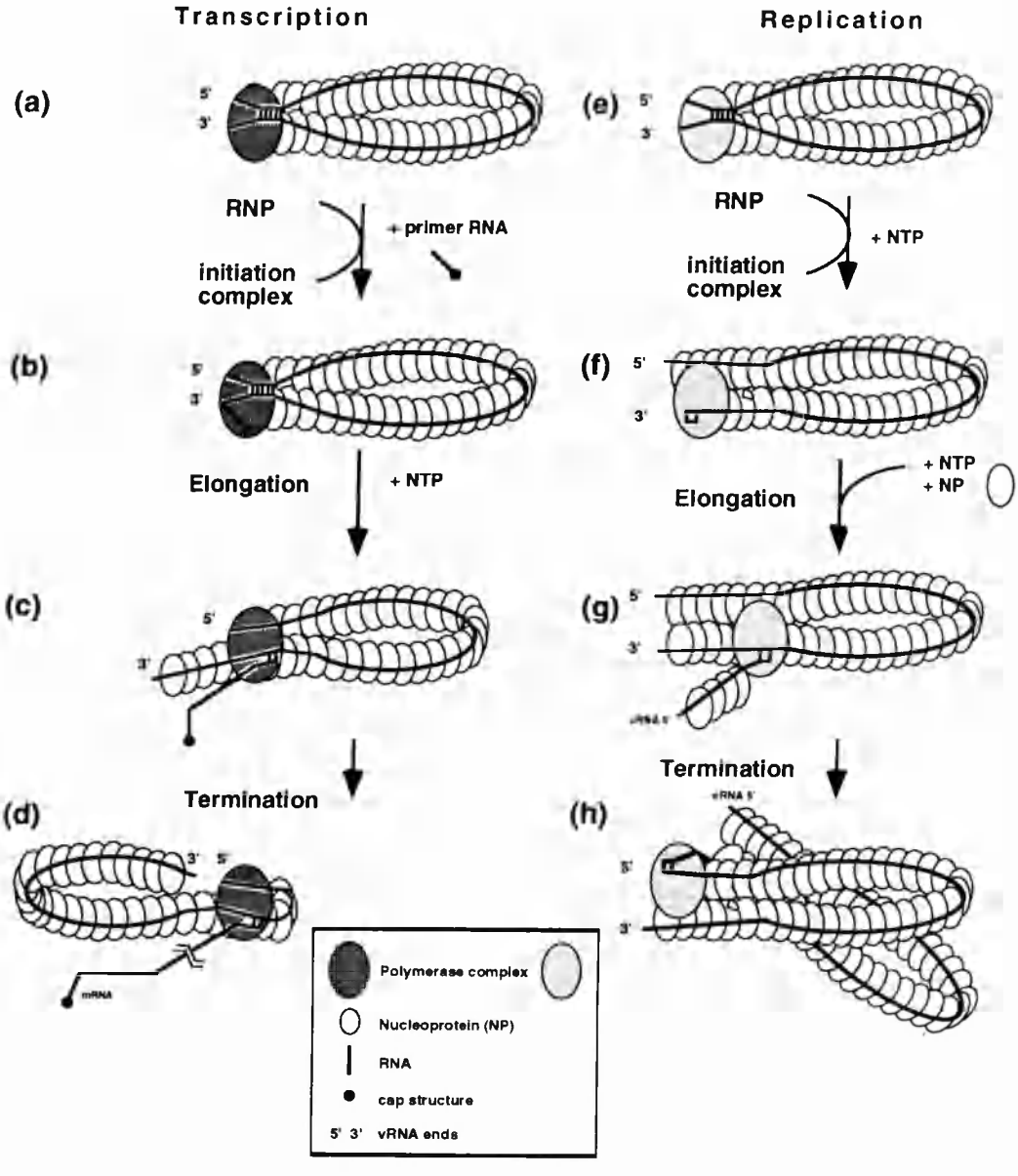


Figure I.18: Possible models of influenza virus transcription and replication combining the present results on panhandle stability with previous results from other laboratories as mentioned in the text. On complete RNPs the vRNA ends are engaged in a ternary complex with the polymerase (a and e). Transcription is initiated with a capped primer RNA (b) and elongation of the primer occurs in the presence of NTPs from the 3' end of the vRNA. During elongation the polymerase might stay bound to the 5' end (c). The vRNA loop on the RNP decreases during this process and transcription termination is imposed by sterical reasons in combination with a "slippery" oligoU stretch (d). A different form of the polymerase performs replication (e), which initiates without a primer at the 3' end of the vRNA and involves an "open" RNA with accessible 5' end (f). Replication is dependent on soluble NP, which binds to the newly synthesized RNA (g). RNA synthesis terminates at the last base of the vRNA 5' end, and complete, cRNA containing RNPs are produced during the process (h).

At present, the available evidence suggests that polymerase binding to the vRNA 5' end is required for transcription initiation from the 3' end, but both ends do not interact in the absence of the polymerase. Similar genome binding patterns have been described for other multi-subunit, RNA-dependent RNA polymerases. The polymerase of Brome Mosaic Virus, a segmented positive strand RNA virus, requires an interaction with an intercistronic region on the genome for initiation of RNA synthesis from the 3' end (Quadt et al., 1995). The replicase of Qβ phage binds to an internal site of the genomic (+)RNA and remains attached there, while initiating (-) strand synthesis from the 3' end. Moreover, the binding pattern to (+) and (-) strand RNAs is different, consistent with different functions of the strands in the replication cycle (Barrera et al., 1993; Schuppli et al., 1994). Poliovirus RNA replication involves polymerase complex formation with both ends of the viral RNA, although in this case the complex formed at the 5' end of (+) RNA has been proposed to catalyze in trans initiation of synthesis from the 3' end of a (-) RNA (Andino et al., 1993; Harris et al., 1994). Finally, the (+) RNA of *Saccharomyces cerevisiae* L-A virus also contains an internal binding site, that binds stronger to the L-A polymerase, than the RNA 3' end, and which is required for in vitro replication (Esteban et al., 1989; Fujimura and Wickner, 1992). This so-called "action at a distance" phenomenon is common in polymerase-enhancer systems to regulate transcription processes and to correctly position the polymerase subunit to the transcription initiation site. The polymerase-promoter interaction itself has to be relatively weak to enable easy promoter clearance after transcription initiation. The strong influenza virus polymerase binding site at the 5' end of the viral RNA assures high specificity recognition of viral RNAs and at the same time brings the polymerase into the vicinity of the low affinity 3' end binding site to start transcription.

The studies on NP interaction with the genomic RNA outline another problem of RNA virus replication, namely the need to release the RNA replicates from the templates in order to make them available for several rounds of RNA synthesis. Many positive strand viruses presumably encode RNA helicases to surmount this problem (Gorbalenya and Koonin, 1993; Lain et al., 1990; Warrener and Collett, 1995), but influenza virus, and most likely other negative strand viruses as well, make use of cooperative single-stranded RNA binding proteins, analogous to the single stranded DNA binding proteins that are cofactors of DNA directed RNA polymerases and DNA polymerases. It has been found that NP binding to RNA removes secondary structures and keeps the RNA single stranded. Influenza virus replication is dependent on soluble NP produced in infected cells, which packages the newly synthesized RNAs during their synthesis. Cooperative NP binding will prevent base-pairing between template and replicate to occur and keep the template available for further rounds of replication. A very similar activity has recently been described during poliovirus replication. The 3D protein of poliovirus displays cooperative single-stranded RNA binding activity during replication (Pata et al., 1995) and it is

thereby able to unwind RNA duplexes of over 1000 base pairs length without the need to hydrolyze ATP (Cho et al., 1993). On the other hand, the viral mRNAs are not dependent on NP to efficiently separate from their complementary template strands. The influenza virus polymerase uses host cell derived, capped RNA primers for transcription initiation, which most likely results in the assembly of cellular cap-binding and hnRNA-binding complexes on the viral mRNA and thereby prevents base-pairing with the template RNA (Izaurralde et al., 1995; Matunis et al., 1993; Piñol-Roma and Dreyfuss, 1992). These examples illustrate that different viruses encounter similar problems during the replication cycle and that apparently there exist only a limited number of concepts to solve them. The negative strand viruses transcribe mRNAs from their genomic RNAs after cell entry, whereas the genomes of the positive strand viruses are already in mRNA sense and can be directly translated in infected cells. This is the major reason for the differences in the genome structure optimized either for virus specific transcription or for translation. The influenza virus RNPs, as packaged into virus particles, are ready to start transcription having the polymerase bound to both vRNA ends and the bases presented for transcription by the nucleoprotein.

Chapter II

Expression of recombinant PA protein

II.1. Introduction

One of the major goals of research on influenza virus replication is to unravel the structure and function of the polymerase complex, which harbours all presently known enzymatic functions required for specific RNA synthesis. In order to better understand the structural basis of these activities, to focus on the biochemistry of certain activities *in vitro* and to exploit them for the development of potential antiviral compounds it is necessary to synthesize and purify the polymerase in reasonable quantities from recombinant sources. Because the polymerase consists of three subunits, encoded by three different viral genes, the problem of expression and purification of the polymerase can either be approached with systems that express all subunits together at the same time, or with systems that express single subunits and make it possible to reconstitute the polymerase complex from the purified precursors afterwards. Recombinant proteins will be needed to map polymerase activities to certain subunits and protein domains and the structures of these domains will help to understand how they function.

II.1.1. Polymerase Subunit Functions

Many laboratories are presently working at mapping known polymerase activities to subunits, trying to establish subunit domains and subunit interaction surfaces. These problems are generally approached using the only system that has convincingly produced functional, recombinant polymerase, namely expression of subunits in mammalian cells. Permanent cell lines derived from murine NIH 3T3 and C127 cells, that express the polymerase subunits, can complement temperature sensitive virus strains (Braam-Markson et al., 1985; Krystal et al., 1986) and are active in influenza virus specific transcription and replication, if the nucleoprotein and a virus-like template RNA are expressed as well (Kimura et al., 1993; Kimura et al., 1992). Similarly, polymerase subunits expressed from recombinant vaccinia or SV40 viruses assemble to active polymerase complexes in a variety of cell types (Huang et al., 1990; de la Luna et al., 1993). These systems have provided the major part of the information concerning the functions of the polymerase subunits. PB1 is most likely the subunit responsible for the actual addition of nucleotides to the growing RNA chains during RNA synthesis. The PB1 protein contains conserved sequence motifs of RNA virus polymerases (Poch et al., 1990), that have been shown by mutational analysis in the recombinant vaccinia virus system to be essential for influenza virus polymerase activity (Biswas and Nayak, 1994). Additionally, the first nucleotide incorporated into transcripts could be UV-crosslinked to PB1, and during elongation PB1 remained at the 3' ends of the growing

mRNAs (Braam et al., 1983). The PB2 subunit can bind to RNA cap structures, and specific anti-PB2 antibodies abolish the endonuclease activity required to produce primer RNAs for viral transcription (Blaas et al., 1982; Blok et al., 1996; Licheng et al., 1995; Ulmanen et al., 1981; Ulmanen et al., 1983). Little is known about the function of PA in the virus life cycle. PA has proven indispensable for transcription and replication in *in vivo* reconstituted systems based on recombinant vaccinia or SV40 viruses (de la Luna et al., 1993; Hagen et al., 1994; Huang et al., 1990). Also, it advances along the viral template RNA and can be crosslinked to the 3' end of the growing mRNA chain during transcription together with the other polymerase subunits, demonstrating that it constitutes an integral part of the viral polymerase complex (Braam et al., 1983). Recent studies, performed with permanent murine cell lines expressing different combinations of influenza virus polymerase subunits, suggest that PB1 alone is sufficient for a basic level of positive strand RNA synthesis and polyadenylation, but PA strongly increases cRNA production and is absolutely essential for vRNA synthesis in this system. PB2 seems to be required only to provide a cap-structure for the mRNAs during transcription (Nakagawa et al., 1995; Nakagawa et al., 1996).

PA might also have an additional function outside the polymerase complex. The nuclear import of PA is delayed with respect to the other subunits and in some systems PA remains partially cytoplasmic (Akkina et al., 1987; Jones et al., 1986; Nieto et al., 1992). A temperature sensitive PA mutant showed deregulated translation of viral proteins (Herget and Scholtissek, 1993). Finally, PA expression appears to be toxic, leads to aberrant nuclear morphology and induces a generalized proteolysis, that reduces the steady-state levels of coexpressed proteins (Sanz-Ezquerro et al., 1995). At the moment it is not yet clear how all these observations can be fitted into an additional function of isolated PA during influenza virus replication. With purified PA protein available it would be possible to test some of the functional hypotheses that have come up from the above mentioned experiments. The identification of an independent PA activity would stand for a novel, influenza virus specific function and eventually a novel antiviral target.

The amounts of recombinant polymerase protein produced in mammalian cells are far too small to attempt purification for structural studies. This chapter summarizes the results from evaluating a number of alternative expression systems for higher level expression and purification of functional influenza virus PA protein.

II.1.2. The PA protein

RNA segment 3 of influenza viruses encodes the PA protein in influenza virus A and B subtypes and the equivalent P3 protein in influenza C viruses. It is presumed that PA/P3 proteins form complexes with the other two polymerase subunits, PB1 and PB2, resulting in active polymerase proteins. However, functional studies have mostly been

done with influenza A viruses. The amino acid sequences of PA/P3 proteins show sequence identities between 28% and 38% among the virus A,B and C subtypes and between 91% and 99% among influenza A virus strains. The extent of conserved amino acid sequences suggest a common ancestor and similar functions for these proteins in the virus subtypes (Akoto-Amanfu et al., 1987; Okazaki et al., 1989; Yamashita et al., 1989). The PA genes of influenza A viruses code for a polypeptide of 716 amino acids with a predicted molecular weight of 82 kD. The sequence alignment of PA proteins from different origins suggests a division of the protein into at least two domains. The N-terminal domain is characterized by the accumulation of mutations, that are specific for virus lineages (human, swine, avian or equine viruses). The N-terminal part has therefore been implicated in the determination of the host range of viruses and possible interactions with cellular factors. The C-terminus is significantly more highly conserved among the PA proteins and is presumed to contain critical structures for the enzymatic activity of the protein (Okazaki et al., 1989). The amino acid sequence of PA shows no striking similarities to other known sequence motifs. There are, however, sequence elements, that resemble the NTP-binding motifs originally described by Walker et al. (figure II.1) (Walker et al., 1982; Yoshida and Amano, 1995).

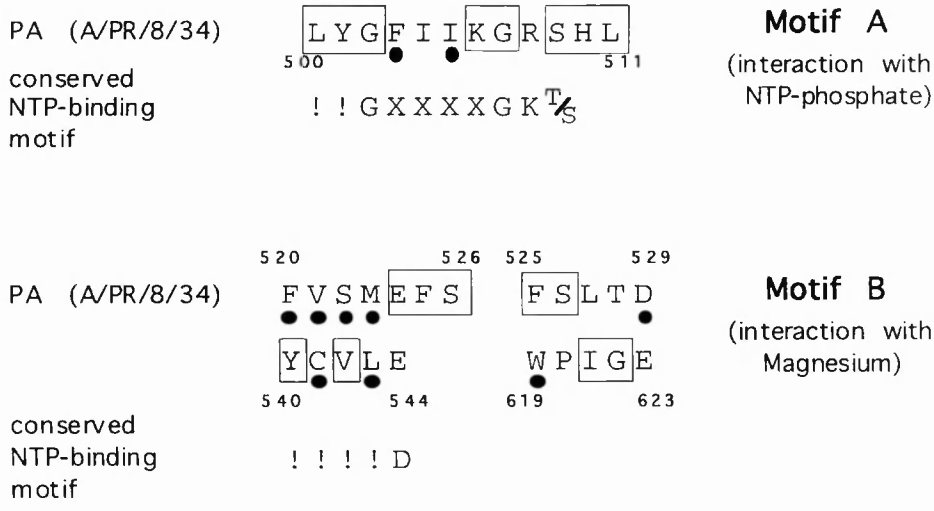


Figure II.1: PA sequence elements, that resemble the Walker NTP-binding motif. Boxed residues are conserved in A, B and C viruses. The residues marked with dots are functionally conserved and have been exchanged only with amino acids of similar hydrophobicity, polarity or charge. "!" stands for hydrophobic residues, X for any amino acid.

All the elements of the possible Walker-like motif are located in a highly conserved region, where no mutations have been introduced during PA evolution in influenza A viruses. This C-terminal region also shows very high conservation among virus subtypes.

The two glycins of the Walker NTP phosphate binding loop are conserved in PA/P3 proteins from all three influenza virus subtypes, but an arginine in motif A is only retained in influenza A and B viruses. The important parameters characterizing the additional elements of the NTP binding motif are not as unique as for motif A, but similar elements are found at similar distances in the PA sequence.

PA from influenza A/PR/8/34 virus has a negative net charge of -13.5 at pH 6.5 and an estimated pI of 5. The amino acid composition shows a very high percentage of glutamic acid (E = 10.75%) and methionine (M = 3.49%), which both are considerably above the average values of 6.16% and 2.27% respectively (Sharp and Cowe, 1991). Secondary structure analysis has been performed with various programs and predicts a high percentage of α-helix and only few β-sheet structures in the protein. Figure II.2 is a graphic representation of the consensus secondary structure as predicted from several algorithms (Chou and Fasman, 1978; Garnier et al., 1978; Geourjohn and Deléage, 1994; Rost and Sander, 1994). The longest α-helical stretches are predicted at the very N-terminus (positions 5-45) and near the C-terminus between amino acids 570 and 610. There is also a large region from residues 140 to 200, which is predicted to be mostly α-helical. These theoretical α-helices are also conserved in the same location in PA/P3 proteins from influenza B and C viruses. These PA protein characteristics have to be considered in the choice of purification conditions and the search for protein folding units by gene fragmentation.

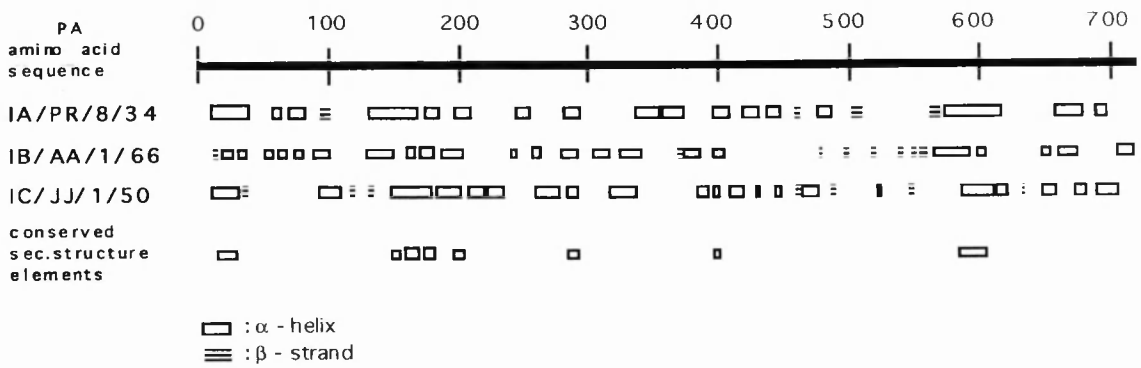


Figure II.2: Secondary structure prediction for PA/P3 proteins from influenza A, B and C viruses. The PA/P3 proteins of the virus strains noted on the left of the picture have been analyzed with several structure prediction programs. The consensus elements have been drawn as rectangulars of different sizes according to the extension of the structures. The bottom line shows the structure elements, that are present in proteins of the three influenza virus subtypes.

II.2. Expression systems used for the production of recombinant PA

Table II.1 gives an overview of the expression systems, that have been used in this study. Although the purification of full-length PA protein was not yet achieved with any of these systems, the experiments provided a number of valuable results concerning possible PA activities in isolated form and suggested possible strategies for the future improvement of PA expression.

Cell Type	organism	purification tag / expression vector	promoter
procaryotic	*E. coli*	N-terminal poly-His-tag pRSET vector (Invitrogen)	T7
		N-terminal Maltose binding protein (MBP) pMAL-c2 (NEB)	*tac*
		N-terminal biotinylated peptide Pinpoint™ sytem (PROMEGA)	*tac*
		N-terminal biotinylated peptide and C-terminal poly-His-tag	*tac*
eucaryotic	SF9, High5 insect cells	no tag recombinant baculovirus	polyhedrin
	Pichia pastoris	no tag recombinant yeast clones	*AOX1*
	Wheat germ extract	no tag pGEM 7Zf+ (PROMEGA)	T7

Table II.1: Compilation of the protein expression systems used to express influenza A virus PA protein

II.2.1 *E.coli* based expression systems

The most widely used expression systems are those using *E.coli* as the host cell. The major reason for this is that in many cases large amounts of recombinant proteins can be produced in a relatively short time due to the rapid doubling time of the bacterium combined with the possibility of using high copy number plasmids for expression. The biology of *E.coli* is well studied and the bacterium can be relatively easily grown and genetically manipulated. A variety of expression systems differing in promoter type, inducibility and protein fusion strategies are commercially available, but their value for the synthesis of recombinant proteins is often largely dependent on specific properties of the protein in question and cannot easily be predicted. Several expression systems have been employed in this study in order to find the optimal combination of parameters for the production of PA protein in *E.coli*. The major problems encountered were protein

insolubility in the pRSET and Pinpoint™ systems and extensive fragmentation of the full-length PA protein in all bacterial systems. A series of N-terminal and C-terminal deletion mutants of PA was screened for the identification of functional domains according to the presumed NTP-binding site at the C-terminus and the nuclear localization sequences at the N-terminus of the protein. Figure II.3 schematically shows the PA fragments, that have been looked at in detail in this study.

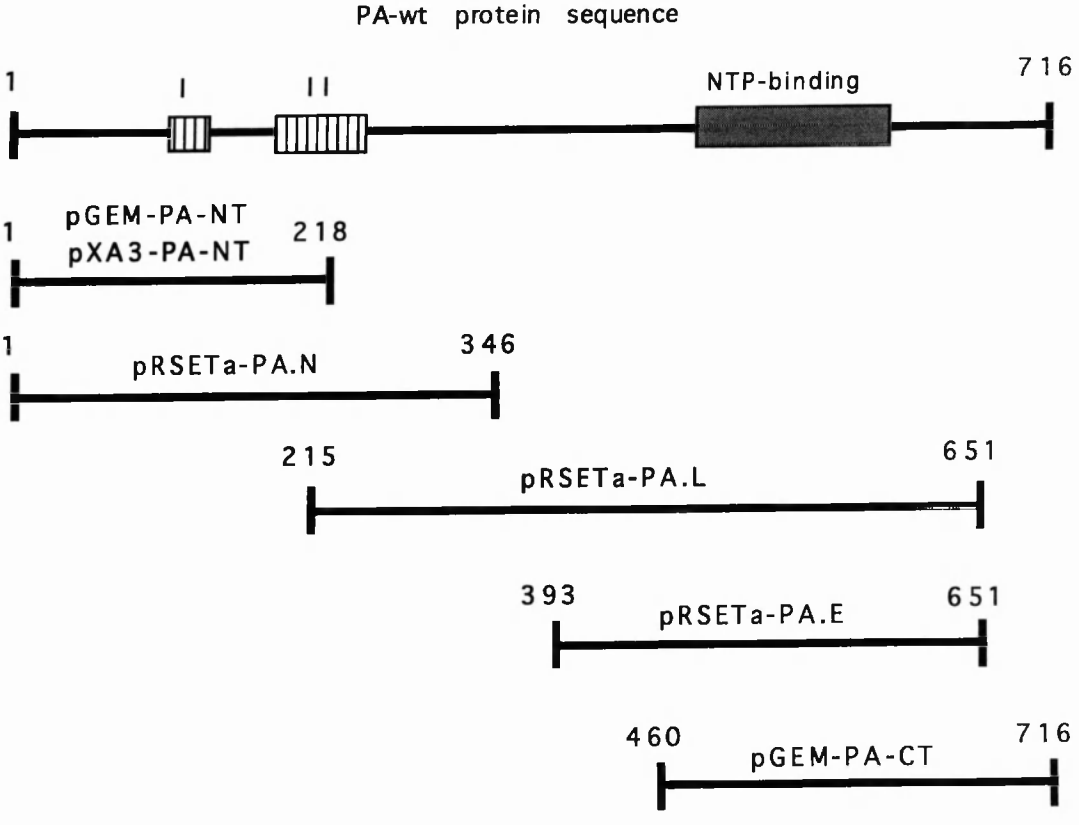


Figure II.3: Schematical overview of the PA deletion mutants, that have been used in this study. The striped boxes on the full-length PA wild-type protein sequence indicate the regions I and II implied in nuclear transport of PA .(Nieto et al., 1992). The shaded box shows the relative position of the sequence motif reminiscient of a NTP binding motif (see figure II.1).

II.2.1.1 Expression of PA from the pRSETa vector

The pRSET vectors promote expression of the recombinant proteins under the control of the T7 promoter and they add a polyhistidine tag to the N-terminus of the protein, that can be removed by treatment with enterokinase. Recombinant *E.coli* host strains are available, that express the T7 polymerase from an IPTG inducible *lacUV5* promoter. The polyhistidine tag provides a possibility for affinity purification of the recombinant protein on Ni-chelate resins. The complete coding sequence of PA and several deletion mutants of PA have been amplified by PCR and cloned into the pRSET vector by Dr. Martin Ford. The pRSET derived constructs, in contrast to the rest of the PA encoding plasmids,

encode additional 22-28 amino acid extensions at the C-termini of the PA fragments. Full-length PA and three PA fragments were expressed in *E.coli* BL21(DE3) cells. The *E.coli* B strains are optimized for protein expression, as they lack the *lon* and *omp* T proteases, thereby reducing protein degradation during cell harvesting and lysis. Induction of PA expression was performed in 500 ml shaker flask cultures at 30° C or 37° C with 100-500 μM IPTG for 2-5 hours. The cells were harvested and lysed by sonication in a buffer containing EDTA and TritonX-100. The expression of protein was analyzed by denaturing gel electrophoresis and Western blot with specific monoclonal antibodies. Under these conditions the majority of PA protein was found insoluble in inclusion bodies, that could be washed with up to 8 M urea without significant loss of recombinant protein. Apart from the respective PA fragments, a large number of smaller bands were also recognized by PA specific antibodies in whole cell extracts and inclusion bodies. This suggested that the PA proteins were either already considerably degraded in the cells or were subject to a deficient processivity during translation. Whereas full-length PA remained largely insoluble even in 6 M guanidinium hydrochloride (Gua-HCl), the PA fragments PA.N, PA.L and PA.E (figure II.3) could be solubilized and purified to homogeneity by Ni-chelate chromatography with a 30-100 mM imidazole gradient. Figure II.4A shows *E.coli* whole cell extracts before (lanes 2 and 5) and after (lanes 3, 4 and 6, 7) induction of PA.L and PA.N, as well as an example of the insoluble fraction (lane 1). Induced bands of the expected sizes are readily visible on this Coomassie stained SDS-PAGE gel and indicated with arrowheads and circles. The purified PA fragments remained soluble after dialysis in renaturation buffer and could be concentrated in Centricon tubes to 2 mg/ml. They were stored in 50% glycerol at -20° C and remained stable for at least one year as judged by SDS-PAGE analysis. About 10-fold higher amounts of recombinant protein were obtained with the 36 kD PA-E fragment from the C-terminus of the PA coding sequence. It was purified, dialyzed in renaturing buffer and stored in the same manner as the above mentioned proteins. Figure II.4B shows the soluble PA-E and PA.L fragments obtained after dialysis in renaturation buffer on a 12 % coomassie stained SDS-PAGE gel, figure II.4C is a Western blot of the same samples. Full-length PA, purified by affinity chromatography partially aggregated during dialysis and partially degraded to a 36kD band, so that the actual full-length band virtually disappeared during dialysis (figure II.4C, arrowhead).

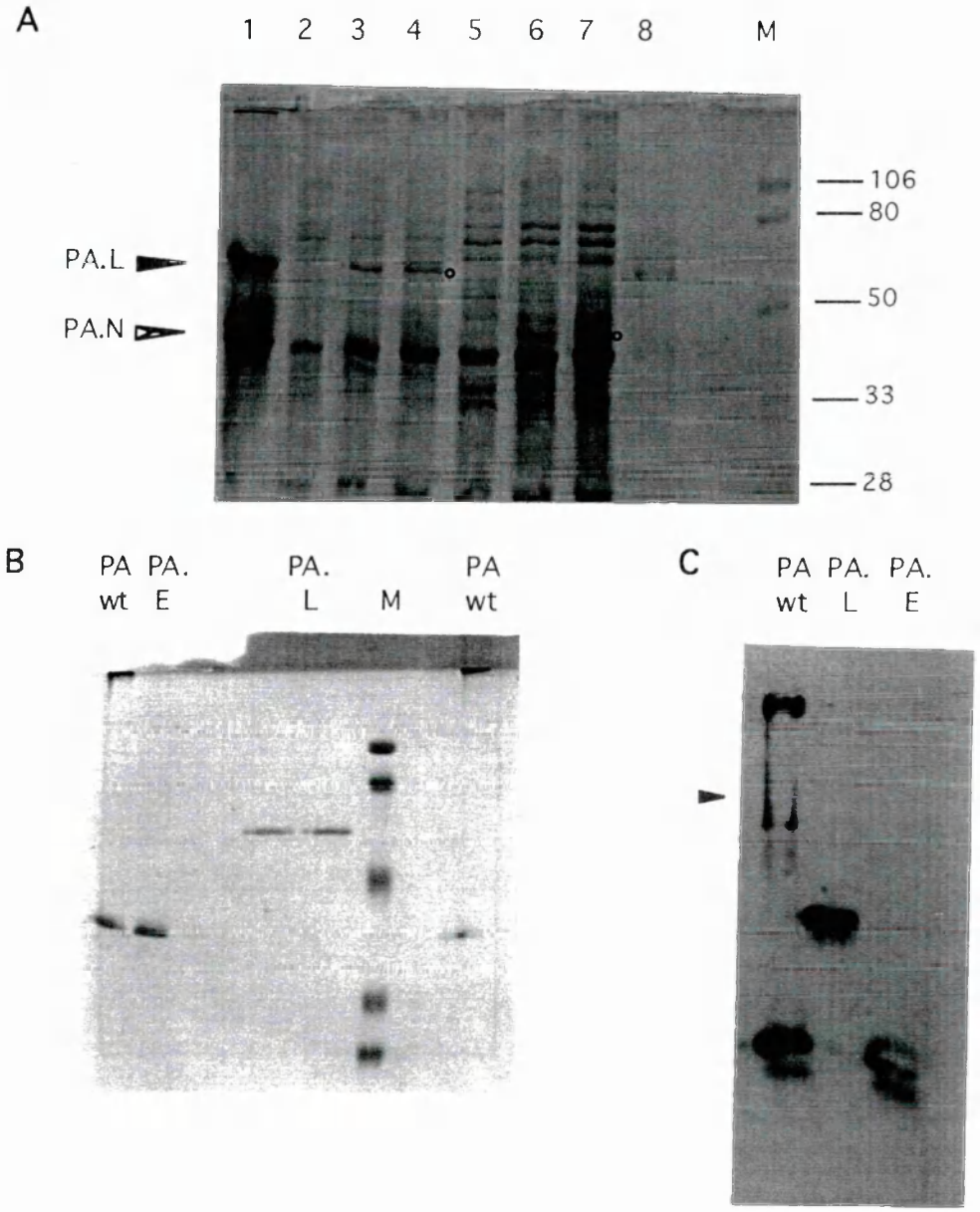


Figure II.4: Expression of PA fragments (see figure II.3). (A): Coomassie-stained, 12% SDS-PAGE gel showing whole cell extracts from E.coli BL21 (DE3) cells transfected with pRSETa-PA.L and pRSETa-PA.N respectively. Protein bands induced after IPTG addition are marked with arrowheads. Lanes **1**, insoluble fraction of PA.L expressing cells; **2**, PA.L expressing cells before induction; **3** and **4**, PA.L expressing cells 2 h and 4 h after induction; **5**, PA.N expressing cells before induction; **6** and **7**, PA.N expressing cells 2 h and 4 h after induction; **8**, insoluble fraction from PA.L expression after washing the pellet with 2 M urea; **M**, BIORAD low range, prestained standards.
(B): Coomassie-stained 12% SDS-PAGE of soluble PA fragments after solubilization in 6M Gua-HCl, nickel affinity chromatography in 8 M urea, dialysis in refolding buffer and concentration in Centricon tubes.
(C): Western blot analysis of the samples shown in (B). The blot was developed with the monoclonal anti-PA antibody L35E12, specific for an epitope near the C-terminus of PA. A small amount of full-length PA protein was detected after dialysis of PA-wt (arrowhead), but most of the immunoreactive material was in large, soluble aggregates or gave rise to a 36 kD breakdown product.

The purified proteins were analyzed for ATPase and RNA binding activity. To this end, the proteins were titrated into an ATPase reaction with radioactive γ-^{32}P-ATP in the presence or absence of *in vitro* transcribed segment 8 vRNA. ATP, ADP and phosphate were then separated by thin-layer chromatogaphy. RNA binding activity was analyzed in filter binding experiments with radioactively labelled segment 5 cRNA or segment 8 vRNA. In both cases no activity was observed with any of the PA proteins. The proteins were also analyzed by circular dichroism spectroscopy (CD) after a buffer change from the renaturation buffer to a phosphate/DTT buffer, pH 7.6. The CD spectra obtained were typical of random coil polypeptides, strongly suggesting that the PA proteins did not refold into functional structures albeit resting perfectly soluble in renaturation buffer (figure II.5).

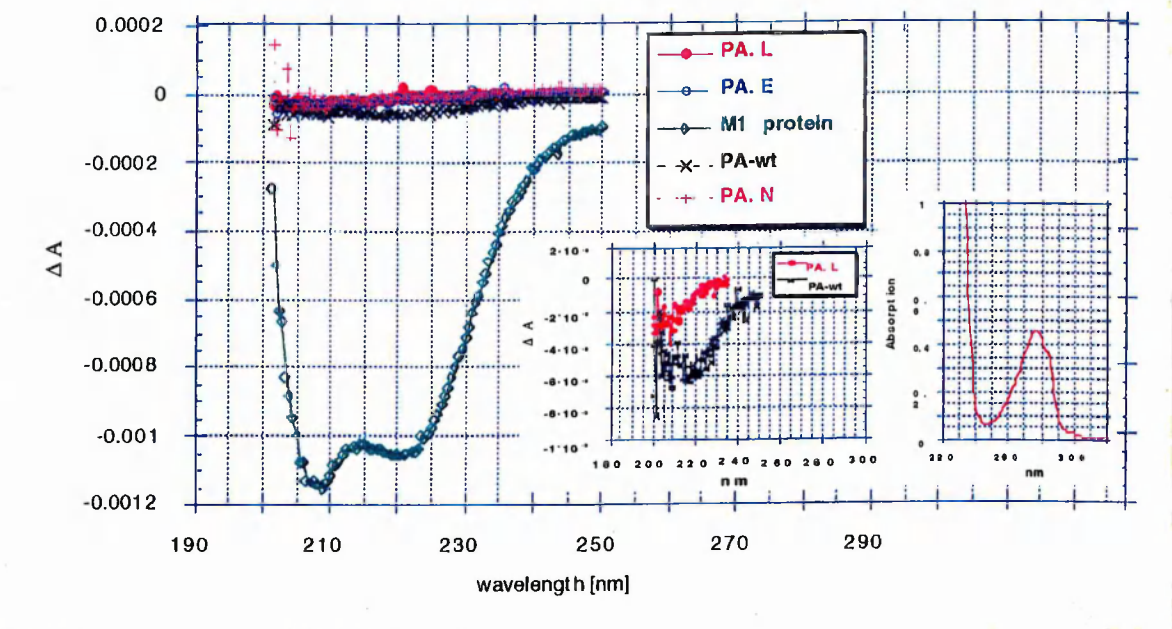


Figure II.5: CD spectroscopy of purified, soluble PA polypeptides compared with purified influenza virus M1 protein, which is predominantly α-helical (provided by Dr. Christine Elster). All proteins were at 0.1 mg/ml. The middle inset shows a magnification of the CD spectra for PA.L and Pa-wt. The apparent minima at 210 and 220 nm may indicate that in the PA-wt sample a small fraction of the protein has adapted a folded conformation. The right inset shows the absorption spectrum of PA.L, which was identical for the other PA fragment and demonstrates the purity and the absence of nucleic acids in the samples.

We did not succeed in refolding any of the PA fragments either by dialyis in different steps or by gel filtration. However, additives like polyethylenglycol, which may help refolding, were not extensively tested. Considering that even the relatively small PA-E fragment did not form stable secondary structures in the renaturation solution, it may be possible that

long-range interactions are important in the formation of the PA tertiary structure. On the other hand, there seemed to be a minute fraction of folded polypeptide in the PA-wt preparation. PA-wt was the only protein preparation that gave rise to multiple bands after dialysis and protein concentration. The spontaneous formation of a 35 kD breakdown product might possibly indicate a fragile spacer region in the protein and the determination of the sequence of this fragment could help to identify a stable C-terminal domain.

II.2.1.2 Pinpoint™ system: expression of PA with a biotinylated peptide tag

The switch from the pRSET to the pinpoint system involves changes in the promoter (T7 to *tac*) and the host cell strain (BL21 to JM109) for protein expression, as well as in the N-terminal sequences of the fusion protein, that might influence translational initiation efficiency. The Pinpoint™ vectors carry a segment encoding a 13kD peptide, which is biotinylated by the biotin ligase of *E.coli* at a single lysine residue. Fusion proteins containing this peptide can be affinity purified on avidin resins. Full-length PA, full-length PA with a C-terminal poly-His tail and a N-terminal fragment of PA were produced by PCR and cloned into the pinpoint pXA3 vector downstream of the *tac* promoter, the biotin-peptide coding sequence and a cleavage site for the endoproteinase factor X. The *tac* promoter is controlled by binding of the *lac* repressor protein and provides inducibility of gene expression by the addition of isopropyl-thio-galactoside (IPTG). Protein expression from the plasmid pXA3-PA after induction with 100 μM IPTG at 30°C or 37°C produced only small amounts of mainly insoluble PA protein. During PA expression, an extra band corresponding to full-length PA, although difficult to see in whole cell extracts, became visible as soon as 2 hours after induction. This band did slightly increase with time. Western blot analysis with the monoclonal antibody L35E12, specific for a C-terminal epitope of PA, showed that, as before in the pRSET system, a large number of smaller, immunoreactive bands were produced together with the 97 kD full-length PA band. An 80 kD PA fragment constituted one of the major bands (figure II.6A, white arrow). The expression of the 40 kD PA N-terminal fragment from plasmid pXA3-PA-NT (see figure II.3) was slightly more efficient and could already be visualized by direct detection with streptavidin-horseradish peroxidase conjugate after blotting onto nitrocellulose, an assay in which the full-length PA was hardly detectable (II.6B). Because most of the recombinant PA protein was insoluble, only very low amounts of PA specific bands could be purified from the soluble fraction on Soft-Link Streptavidin resin. The alternative expression plasmid pXA3-His-PA contains a C-terminal polyhistidine tag in addition to the N-terminal biotinylated peptide. The plan was to separate full-length PA specifically from the smaller PA fragments in two successive affinity purification steps. However, the amount of soluble PA protein obtained after the streptavidin column was not sufficient as starting material for a following Ni-chelate chromatography.

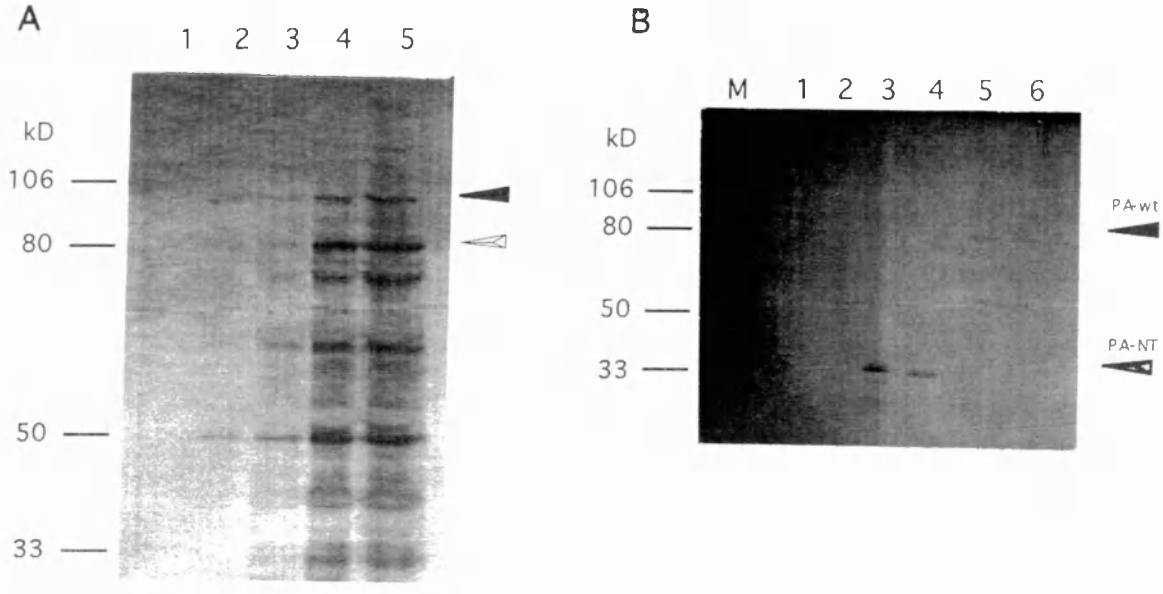


Figure II.6: Western blot analysis of PA expression in the pinpoint system. (A): Western blot of a time course of PA expression in JM109 cells, developed with L35E12 C-terminal anti-PA antibody. Lanes **1**, whole cell extract of PA-NT 3 hours after induction (negative control, showing the specificity of the antibody); **2**, whole cell extract of PA-wt before induction; **3-5**, PA-wt cell extracts 1, 3 and 5 hours after induction. The positions of molecular weight markers are shown on the left. The black arrowhead indicates the band that corresponds in size to the expected 97 kD full-length biotinylated PA fusion protein, the white arrowhead a major 80 kD PA fragment (see text).
(B): Western blot directly developed with peroxidase-coupled streptavidin. Lanes **1**, control cell extract 3 h after induction; **2**, control cells before induction; **3**, PA-NT expressing cells 3 h after induction; **4**, PA-NT cells before induction; **5**, PA-wt expressing cells 5 h after induction; **6**, PA-wt cells 3 h after induction. The arrowheads on the right indicate the positions of crossreactive bands.

Low level protein expression as well as protein fragmentation in *E.coli* might result from limiting amounts of so-called rare tRNAs, because proteins that are highly expressed in *E.coli* show a codon bias towards 22 of the 61 coding triplets. The sequences of eucaryotic genes are known to sometimes deviate considerably from this preferred codon usage in bacteria. In particular the arginine encoding triplets AGA and AGG are common in eucaryotic genes, but they are the least used codons in *E.coli* (Mattes, 1993). It has also been reported that the limiting levels of the $tRNA_{arg4}$, that translates the AGA codons, can cause +1 ribosomal frameshifts at tandem AGA and AGG codons .(Spanjaard et al., 1990). The *argU* gene, encoding the $tRNA_{arg4}$, has been cloned into an helper plasmid and successfully used for improving the expression of various human and plant genes in *E.coli* (Schenk et al., 1995). Because 32 out of 40 arginines in the PA protein are encoded

by these rare codons and tandem rare arginine codons occur twice in the PA sequence, the effect of a coexpression of the the $tRNA_{arg4}$ was analyzed in the pinpoint system. The pACYC177 based tRNA expression vector pUBS520 was obtained from Dr. Ralf Mattes and the tRNA coexpressed in JM109 cells together with wild-type PA. Figure II.7 illustrates the effect of this coexpression with 100 μM IPTG induction at 30°C for 3 hours. A strong band is induced in the cells coexpressing pUBS520 (lane 5, black arrow), but not with pXA3-PA alone (lane 2). The corresponding Western blot (figure II.7B) shows that the coexpression of the rare arginine codon tRNA significantly increases the amount of PA specific protein bands (compare lanes 1-3 with lanes 4-6). The strongly induced band in figure II.7, lane 5 migrates at about 80 kD in the SDS gel and might therefore constitute the major PA fragment, that was mentioned above. This 80 kD polypeptide completely remains in the insoluble fraction (lane 6) and is absent in the fraction loaded onto the streptavidin column (lane 7) so, consequently, no protein is eluted from the column (lanes 8-11). A major 80 kD band was also seen on the Western blot developed with the L35E12 monoclonal antibody (figure II.7B, black arrowhead) together with the full-length PA band (sprinkled arrowhead). This 80 kD polypeptide might either constitute a relatively stable intermediate of proteolysis or be derived from translation termination at a signal other than a frameshift induced by consecutive, rare arginine codons in the PA sequence. These results demonstrate that the coexpression of the tRNA expressing plasmid pUBS520 can significantly increase the total yield of recombinant PA protein in *E.coli*. Another striking difference between PA expression in the absence (lanes 1-3) or presence (lanes 4-6) of overexpressed "rare" tRNA was the production of a strong 32 kD band only when tRNA was limiting (figure II.7A, white arrow). A band of this size is expected to be produced from a theoretical +1 frameshift at the tandem rare argine codon motif at position 369 in the PA gene sequence. Taken together, the large number of rare arginine codons does indeed appear to significantly lower PA yield in this bacterial system, but overexpression of the corresponding tRNA was not sufficient to prevent PA fragmentation. In addition, the presence of the $tRNA_{arg4}$ did not influence the solubility of the PA protein in this system. Further evidence for a role of rare codons in reducing PA yield were obtained from the maltose binding protein system (see II.2.1.3 and discussion II.3).

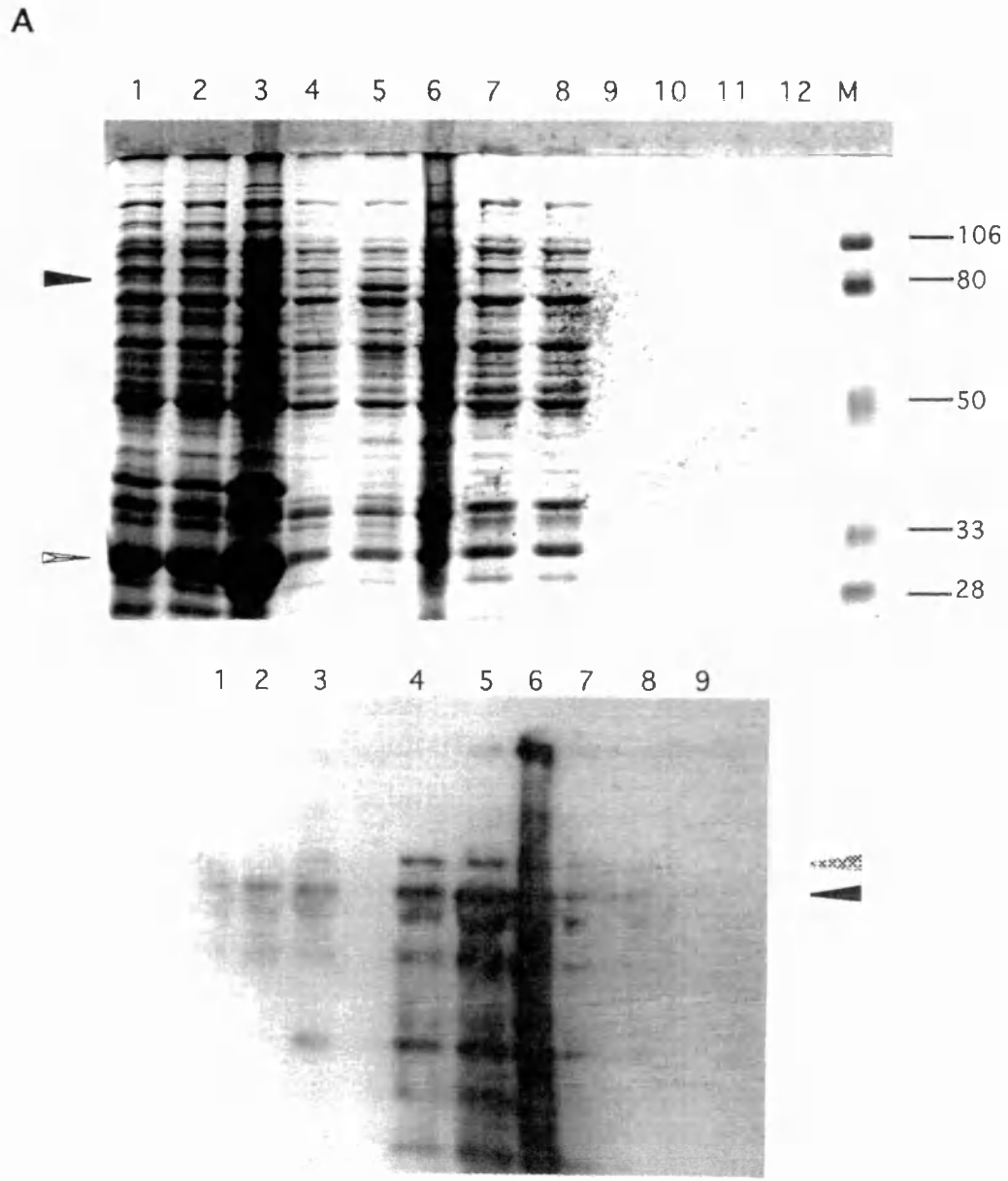


Figure II.7: Effect of the coexpression of PA and a rare arginine codon tRNA in JM109 cells. Samples without, lanes 1-3, or with, lanes 4-12, tRNA coexpression.
(A), Coomassie-stained 12 % SDS-PAGE; the black arrowhead indicates an overexpressed 80 kD polypeptide in lane 5, the white arrowhead a 32 kD band only present in lanes 1-3; **(B)**, Western blot with L35E12 C-terminal anti-PA antibody. The sprinkled arrowhead indicates full-length PA, the black arrowhead a 80 kD PA fragment.
Lanes **1** and **4**, cell extract before induction; **2** and **5**, cell extract after induction; **3** and **6**, insoluble fraction; **7**, soluble fraction; **8**, flow-through of streptavidin column; **9-12**, elution from streptavidin column; **M**, BIORAD low range, prestained standards.

II.2.1.3 The Maltose-binding-protein system (MBP)

In the MBP system the relatively large maltose binding protein (43 kD) is fused to the N-terminus of the recombinant protein sequence and the fusion protein is expressed from the IPTG-inducible *tac* promoter (see II.2.1.2). The use of an *E.coli* derived protein at the N-terminus of a fusion protein ensures optimally adapted transcription and translation initiation in the bacterial host cells. The MBP is well expressed, highly soluble in *E.coli* and can increase the expression and the solubility of attached polypeptides. In addition, the specific binding of MBP to maltose resins can be exploited for affinity purification of the fusion proteins. There is a polyasparagine spacer and a factor X recognition sequence between the MBP and the recombinant protein in order to insulate the MBP domain and to allow the specific disconnection of MBP respectively. The pMAL-PAM2 plasmid contains the complete coding sequence of PA downstream of the factor X cleavage site, pMAL-PAB2.7 encodes six additional amino acids between the cleavage site and the PA start codon (ISEFGS). Both plasmids were provided by Dr. Martin Ford.

The use of the maltose binding protein (MBP) fused to the N-terminus of PA led to a significant increase in the solubility of the PA protein, yet significant fragmentation was again reproduced. Figure II.8A shows a typical purification profile on an amylose resin. Although an additional band of the expected size is visible in whole cell extracts after induction and is not present in the flow-through fraction, this band is not concentrated on the amylose column and most of the eluted proteins are of smaller size. The majority of these eluted bands appear to crossreact with monoclonal antibodies raised against an N-terminal fragment of PA (figure II.9B, right panel). However, only a small number of the bands contain the epitope for the monoclonal antibody L35E12, that was raised against a C-terminal fragment of PA (II.8B). This remarkable difference might be caused by a protection from proteolysis of the PA N-terminus by the MBP moiety. On the other hand, inefficient translation with many stops, once the coding sequence of the MBP has been passed by the ribosomes, would also result in this biased pattern of fragmentation. A number of different cell lines were compared for PA expression, but all produced the same pattern of fragmentation (figure II.8B). Strikingly, full-length PA was produced in largely unfragmented form in the absence of IPTG in uninduced cell cultures. This leaky PA production stems from non-repressed *tac* promoters due to limiting concentrations of the *lac* repressor. Accordingly, much less PA was produced in uninduced JM 109 cells, which overexpress the *lac* repressor. The strong increase in fragmentation upon activation of large numbers of *tac* promoters may reflect the increased limitation of translation by rare tRNAs. However, as was seen in the pXA3 system, the coexpression of pUBS520 (see II.2.1.2) did not significantly reduce the number of PA fragments produced, but rather led to an increase in total protein yield (not shown). This could either imply that at least some of the PA fragments in figure II.8B are derived from

proteolysis, or that rare codons other than arginine codons are involved (see table II.3 in discussion II.3).

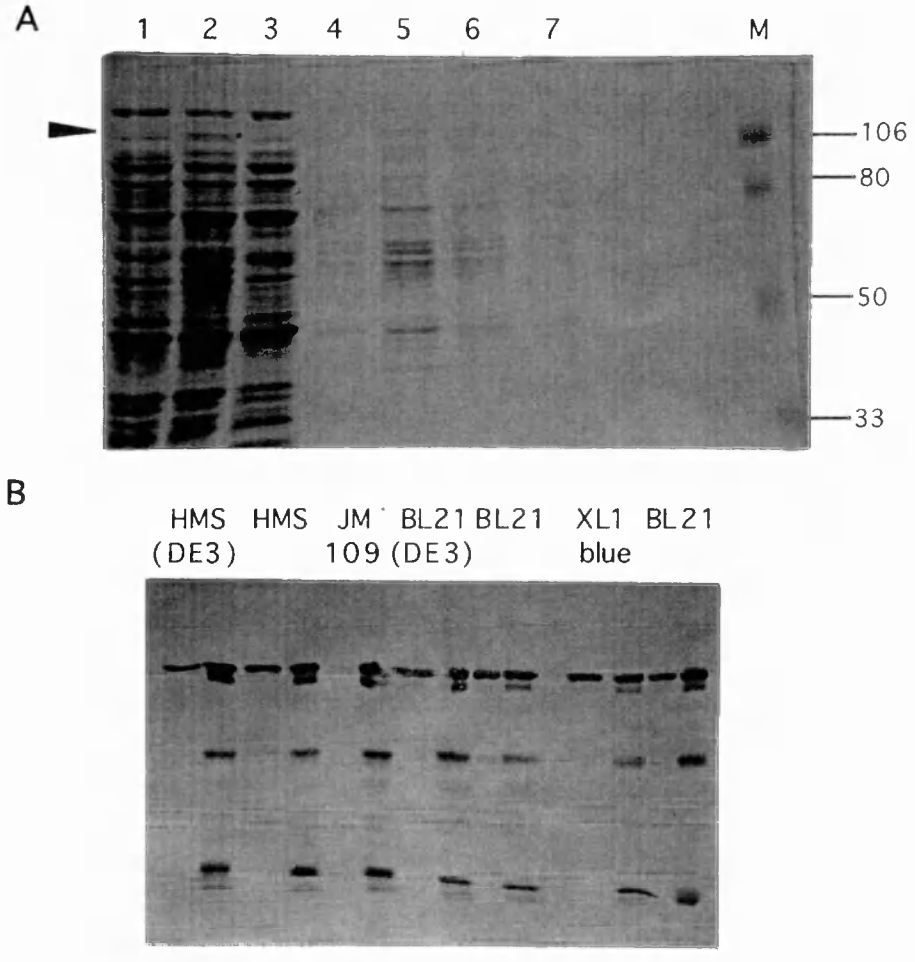


Figure II.8: Expression of PA as a MBP-fusion protein. (A), Coomassie stained, 12 % SDS-PAGE gel. PA was expressed in JM109 cells from plasmid pMAL-PAM2. An additional band of the expected size (124 kD) is induced by IPTG (lane 2) and indicated by the arrowhead and a circle. Lanes 1, cell extract before induction; 2, cell extract 3 h after induction at 37°C; 3, flow-through fraction of an amylose column; 4-7, elution of bound protein from an amylose column; M, BIORAD low range, prestained standards.
(B), Comparison of different *E.coli* strains for MBP-PA expression. Western blot of whole cell extracts developed with L35E12 anti-PA antibody. Pairs of uninduced (left) and 3 h induced (right) whole cell extracts have been loaded onto the gel. The corresponding *E.coli* strain is indicated above these pairs.

The affinity purified MBP-PA fusion protein preparation was analyzed for ATP binding and ATPase activities. ATP binding was performed with periodate oxidized ATP, that crosslinks to proteins by forming covalent bonds with exposed amino groups. Specific nucleotide binding sites can be identified in the presence of excess unlabelled ATP, which efficiently protects these sites from being crosslinked (Clertant and Cuzin, 1982; Jindal et al., 1994). The MBP-PA fusion protein was incubated with oxidized ATP in the presence of different nucleotides, the resulting Schiff-base between oxidized ATP and the protein was reduced with cyanoborohydride, the crosslinked proteins were separated by SDS-PAGE, blotted onto nitrocellulose membranes and analyzed by autoradiography (figure II.9). Two major bands in the PA preparation are reproducibly labelled with oxidized ATP. They are polypeptides of about 85 kD and 80 kD respectively. The labelling is specific for ATP, because only ATP, but not UTP, CTP or GTP can compete for the crosslinking by oxidized ATP. In addition, specific labelling of proteins after amylose affinity purification was only observed with oxidized ATP, but not with other oxidized NTPs (not shown). These results indicate that either PA is an ATP binding protein, or another ATP binding protein is copurifying with PA on the amylose resin. In an attempt to identify the ATP binding polypeptides, the affinity purified PA protein preparations were preincubated with anti-PA antibodies. Neither the monoclonal antibody L35E12, raised against the C-terminus of PA, nor a pool of monoclonal antibodies specific for the N-terminus of PA could inhibit ATP labelling of the 85 kD and 80 kD bands (figure II.9, lanes 7, 8). We never detected labelling with ATP of the full-length MBP-PA fusion protein (127 kD). The same nitrocellulose membrane, that had been used for autoradiography was subsequently probed with a pool of monoclonal antibodies to the N-terminus of PA (figure II.9B, right panel). As mentioned above, the N-specific antibodies detected a large number of PA derived protein fragments, whereas they did not crossreact with any bacterial proteins. The 85 kD and 80 kD ATP binding polypeptides comigrate with two major bands on the corresponding Western blot, leaving it possible that the ATP labelled proteins are PA fragments derived from the MBP-PA fusion protein. However, if the ATP binding polypeptides are to comprise both MBP, for binding to the amylose column, and the PA N-terminus, for crossreaction with anti-PA-N antibody, they are unlikely to extend to the presumed ATP-binding motif in the C-terminal half of PA. Accordingly, the analysis of blotted ATP-binding reactions with PA C-terminus specific L35E12 antibody failed to establish the presence of the C-terminal epitope on the ATP-binding polypeptides.

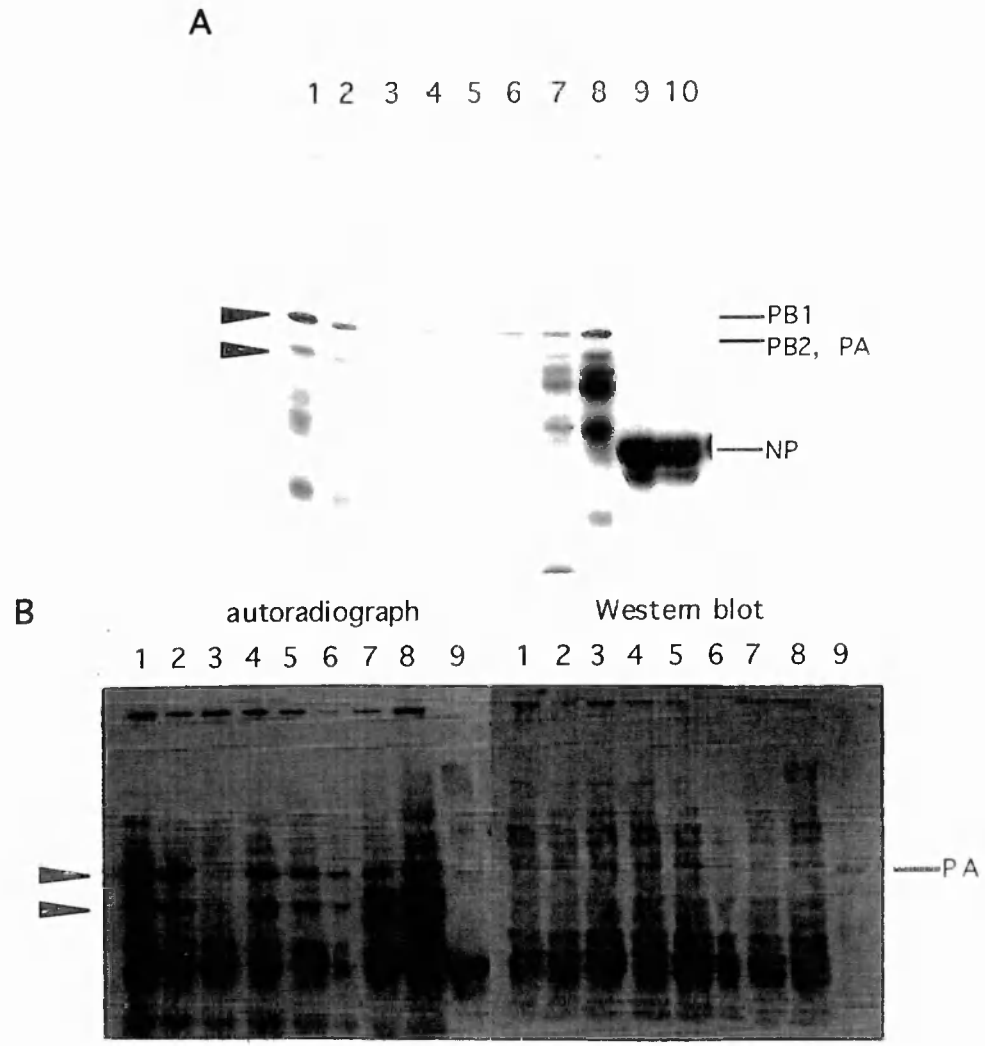


Figure II.9: ATP binding of affinity purified MBP-PA protein preparations. (A), autoradiogram of a 12% SDS-PAGE gel showing ATP crosslinking to a MBP-PA fusion protein preparation. Essentially all polypeptides in the preparation are radioactively labeled with oxidized ATP. Binding specificity is shown by specific competition with unlabeled nucleotides. Lanes, **1** and **2**, 1 μg and 0.6 μg MBP-PA incubated with 0.25 μM ox-ATP; **3-6**, binding reactions with 0.6 μg MBP-PA in the presence of 5 mM unlabeled ATP (**3**), UTP (**4**), GTP (**5**) and CTP (**6**); **7** and **8**, binding reactions with 0.6 μg MBP-PA after preincubation with 8 μl L35E12 antibody (**7**) or N-antibody (**8**); **9**, binding reaction with 0.25 μg influenza virus RNPs. The positions of RNP proteins are marked on the right. The arrowheads indicate the positions of two specific ATP binding polypeptides. (B), Autoradiogram (left) and immunodetection of PA specific bands with monoclonal anti-PA N-terminus antibodies (right) after blotting of the binding reactions shown in (A) onto a nitrocellulose membrane.

Consistent with the ATP binding results, the MBP-PA preparations showed significant ATPase activity *in vitro*. As described above (II.2.1.1), ATP hydrolysis was followed by separation of the products on TLC plates. As shown on figure II.10, the PA preparation produced free phosphate from γ^{32}P-ATP (lane 4). This reaction was independent of added RNA (lane 5). Lane 6 is a control reaction performed with a fraction eluted from the same amylose column after the passage of an *E.coli* cell extract containing the pGEM7Zf+

plasmid. These cells do not produce PA protein, were lysed in the same way as the MBP-PA expressing cells and the cell extract passed through an amylose resin. In the fractions corresponding to the PA elution fractions, no proteins were visible on Coomassie stained gels, and no ATPase activity was detected. These results indicate that the present purification of a polypeptide with ATPase activity on the amylose resin was dependent on MBP-PA expression.

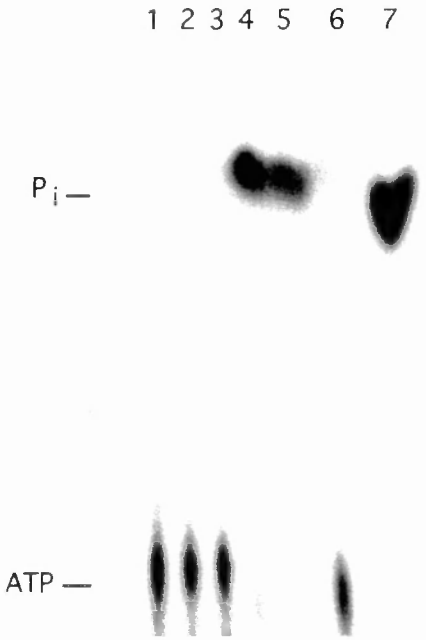


Figure II.10: ATPase activity in amylose purified MBP-PA protein preparations. 0.6 μg MBP-PA was incubated with [γ-^{32}P]-ATP in ATPase buffer. 30 ng aliquots were analyzed by thin-layer chromatography in 1 M LiCl, 0.5 M formic acid. Lanes **1**, ATP control incubation in the absence of protein; **2** and **3** early elution fraction from amylose resin without (**2**) or with (**3**) 1μg NS vRNA; **4** and **5**, PA elution peak from amylose resin without (**4**) or with (**5**) 1 μg NS vRNA; **6**, fraction corresponding to PA elution peak, obtained by passing control cell extract through the same amylose column; **7**, phosphate standard, derived from incubation of [γ-^{32}P]-ATP with alkaline phosphatase.

II.2.2 Expression of PA from recombinant baculoviruses

In order to get around the problems of unfavourable codon bias of the PA gene for expression in *E.coli* and protein fragmentation, eucaryotic protein overexpression systems were tested for PA protein production. In the baculovirus system recombinant proteins are expressed from the polyhedrin promoter during the late stage of a viral infection of insect cells. The system relies on the construction of recombinant baculoviruses through homologous recombination of wild-type virus and special transfer vectors (O'Reilly et al., 1992). The polyhedrin promoter is especially strong, because large amounts of polyhedrin protein are required in the course of a normal infection cycle to package virus particles into inclusion bodies, called polyhedra. These polyhedrin based

inclusion bodies are ingested by insect larvae and start a new infection cycle. The polyhedrin protein is, however, not essential to maintain a baculovirus infection in cultured insect cells, which means that the strong, late polyhedrin promoter can be used to express foreign proteins. The PA protein has been repeatedly reported to possess considerable toxicity in mammalian cells. It was therefore considered to be a major advantage of the late expression in the baculovirus system that recombinant protein synthesis starts only after the maturation of new, infectious virus particles. This ensures that even the expression of cytotoxic proteins will not prevent virus replication. The *AcMNPV* (*Autographa californica* multinuclear polyhedrosis virus) derived, recombinant baculovirus *BacPAK6* (Kitts and Possee, 1993) and transfer vectors provided by Dr. Robert Krug (St.Angelo et al., 1987) were used to produce recombinant baculoviruses containing the coding sequences of all three influenza virus polymerase subunits. The recombinant, plaque purified viruses were then analyzed by DNA dot blot assays for the incorporation of the polymerase coding sequences. Dot blot positive viruses were obtained for the three influenza virus polymerase subunit genes.

II.2.2.1 Time course of expression

The recombinant baculoviruses expressed the influenza virus polymerase subunits only at low levels, and no additional bands were observed on Coomassie-stained acrylamide gels comparing total insect cell extracts after infection with recombinant and control viruses. However, all three subunits could be identified on Western blots. Figure II.11 shows a time course of expression for the different proteins in whole cell extracts. PA was detectable at 32 hours *p.i.* and reached maximal level at 40 hours *p.i.* Significant amounts of degradation products were detected at 72 hours *p.i.* and they increased in intensity thereafter. In contrast to the bacterial systems, PA is mainly produced in full-length form albeit at very low levels. Similarly, the PB1 protein was expressed at low levels and did not accumulate to larger amounts during the time course. A β-galactosidase expressing baculovirus was used as a control in these experiments. Because the polyclonal antisera against PB1 and PB2 had been raised in rabbits against β-galactosidase fusion proteins, the control extracts were strongly labelled by the antisera. It was not determined which of the lower molecular weight bands in the PB1 time course were PB1 specific degradation products and which were unspecific crossreactions of the antiserum. In contrast to the other two subunits, PB2 appeared to slightly accumulate with time. However, in all cases the recombinant protein levels remained below 0.5% of total cellular protein.

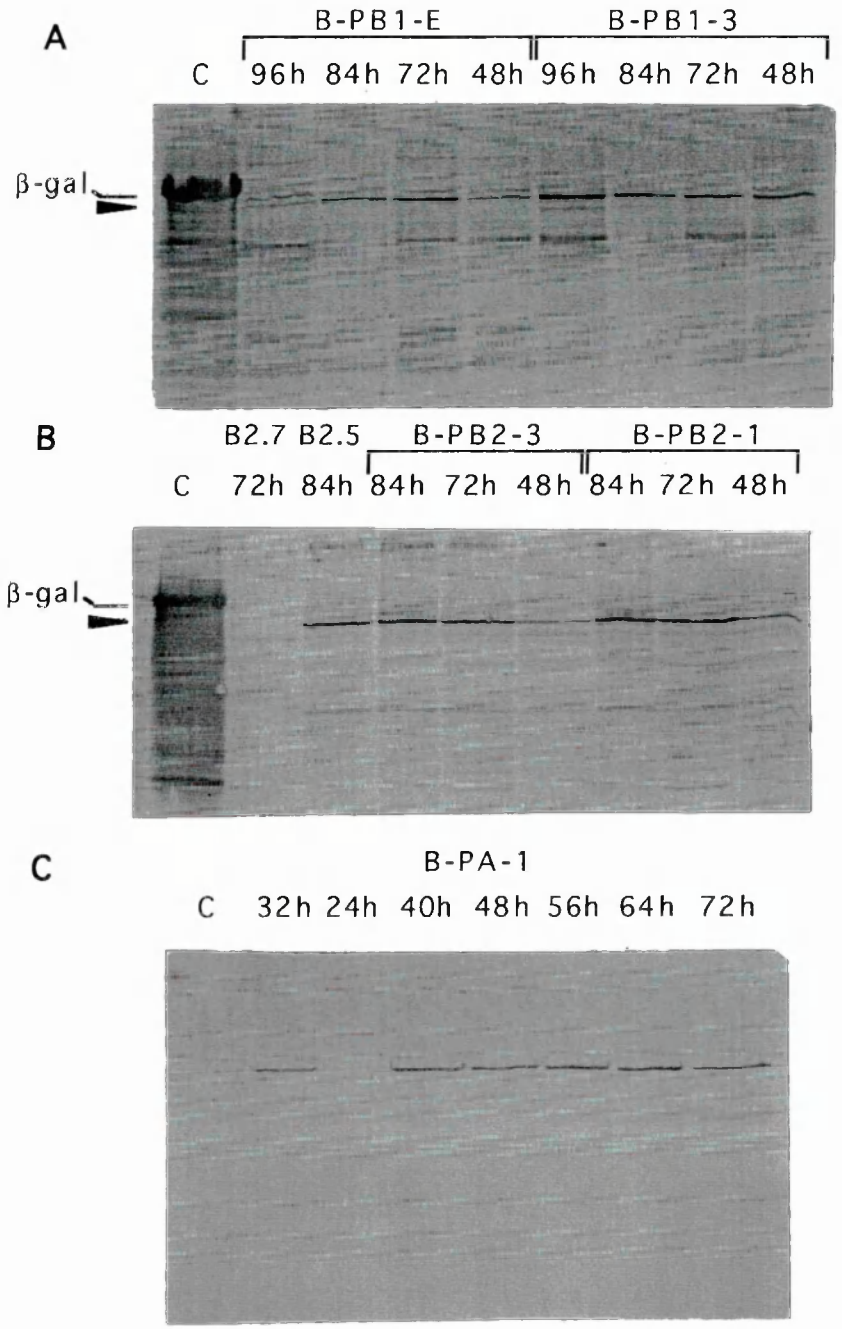


Figure II.11: Western blot analysis of protein expression in baculovirus infected SF9 cells. $3 \cdot 10^7$ cells have been infected with 10 pfu/cell of recombinant viruses expressing PB1 (A), PB2 (B) or PA (C), as indicated above each picture. Two different clones are shown for PB1 expressing viruses, four clones for PB2 expressing viruses and one for PA. Aliquots were taken at different time points after infection and total cell extracts separated by 12% SDS-PAGE. The blots were developed with polyclonal antisera against PB1- and PB2-β-galactosidase fusion proteins, provided by Dr. Paul Digard, and L35E12 monoclonal antibody against PA. Lane "C" contains cell extracts from SF9 cells infected with a β-galactosidase expressing baculovirus. The position of β-gal is marked on the left (116 kD), and arrowheads denote the presumed polymerase specific bands.

II.2.2.2 Cell lysis and clarification

Up to 50% of PA remained insoluble, when the cells had been lysed with a standard buffer containing between 0 and 1% NP40 and sonication. In contrast, most of PA was found soluble in the supernatant of cytoplasmic extracts prepared by incubation of insect cells in a hypotonic lysis buffer and 0.5 % NP40 followed by a 20000xg centrifugation. This result suggested that PA can become attached to insoluble macromolecules during the preparation of whole cell extracts or may become denatured by detergent treatment and sonication. In addition, this confirmed previous observations by others, who had reported that PA was not transported into the nucleus of insect cells in contrast to mammalian cells (Kobayashi et al., 1992). It is also an illustration of the idea that the nuclear localization signal of PA is more complex than that of other influenza virus encoded nuclear proteins (PB1, PB2, NP), that are transported to the nucleus even in insect cells (Galarza et al., 1992; Kobayashi et al., 1992). The possibility to prepare a cytoplasmic extract containing most of the recombinant PA provided an important first step in PA purification. In particular, the contamination of PA preparations with large amounts of nucleic acids was avoided. A major problem of cell extract preparation was to avoid extensive PA degradation. Whereas most of the protein was produced in full-length form in the cells, it very rapidly became degraded during cell harvesting, presumably due to the disruption of cellular lysosomes. A significant amount of PA was also detected in the supernatant of insect cell cultures, probably originating from cell leakage during the harvesting procedure. The PA from this source was virtually free of degradation products as judged by Western blot and the large volume promised considerable amounts of recombinant protein to be present in this fraction. Taken together a very gentle cell lysis with minimal disruption of lysosomes may be crucial for an optimal harvesting procedure of PA expressing insect cells. PA protein became relatively stable after one purification step on an anion exchange column (see below).

II.2.2.3 Anion exchange chromatography of PA

The goal of the first purification step was to concentrate the PA protein while removing as much of the major contaminants as possible. The relatively low pI of 5 was taken into consideration for MonoQ anion exchange chromatography, which was performed in Bis-Tris buffers, pH 6, 50 mM NaCl. Under optimal conditions about 75% of the loaded proteins were separated from the PA (figure II.12), which was mainly eluted at about 300 mM NaCl. The PA peak fractions contained full-length PA protein, stable for at least one year when stored in glycerol at -20°C as judged by Western blot. However, although a distinct PA elution peak was evident (fractions 48-52), PA was not significantly concentrated in the eluted fractions compared to the starting material ("load"), and the PA

protein remained a minor band on Coomassie stained gels (not shown). PA could also be bound to a MonoQ column from cell culture supernatant with a similar purification efficiency (figure II.13). About 97 % of total protein, in this case mainly serum albumin, were separated by the ion exchange procedure. Again, a polypeptide band of full-length PA size accumulated in the column fractions of the PA peak (figure II.13B). The same Western blot shows PA also in flow-through and washing fractions (1, 10, 15, 31), because the MonoQ column was voluntarily used at maximal binding capacity in order to obtain maximally concentrated PA elution fractions, that might have facilitated further chromatographic steps. However, the PA protein was never observed as a major protein in the elution fractions, due to the low abundance in the starting material. Because of the low PA concentration, further purification procedures were difficult to assess. PA was found to bind to Reactive Red and Reactive Green dyes and could be eluted from such dye columns with salt gradients. However, PA binding to the dyes was never quantitative (30-50% binding) and the resolution was low, presumably because the dye columns mainly functioned as anion exchangers in this case and were therefore not well suited to resolve proteins, that had previously bound to a MonoQ column. As a result, the reactive dye columns separated only another 50-70% of the loaded protein from the PA elution peak fractions.

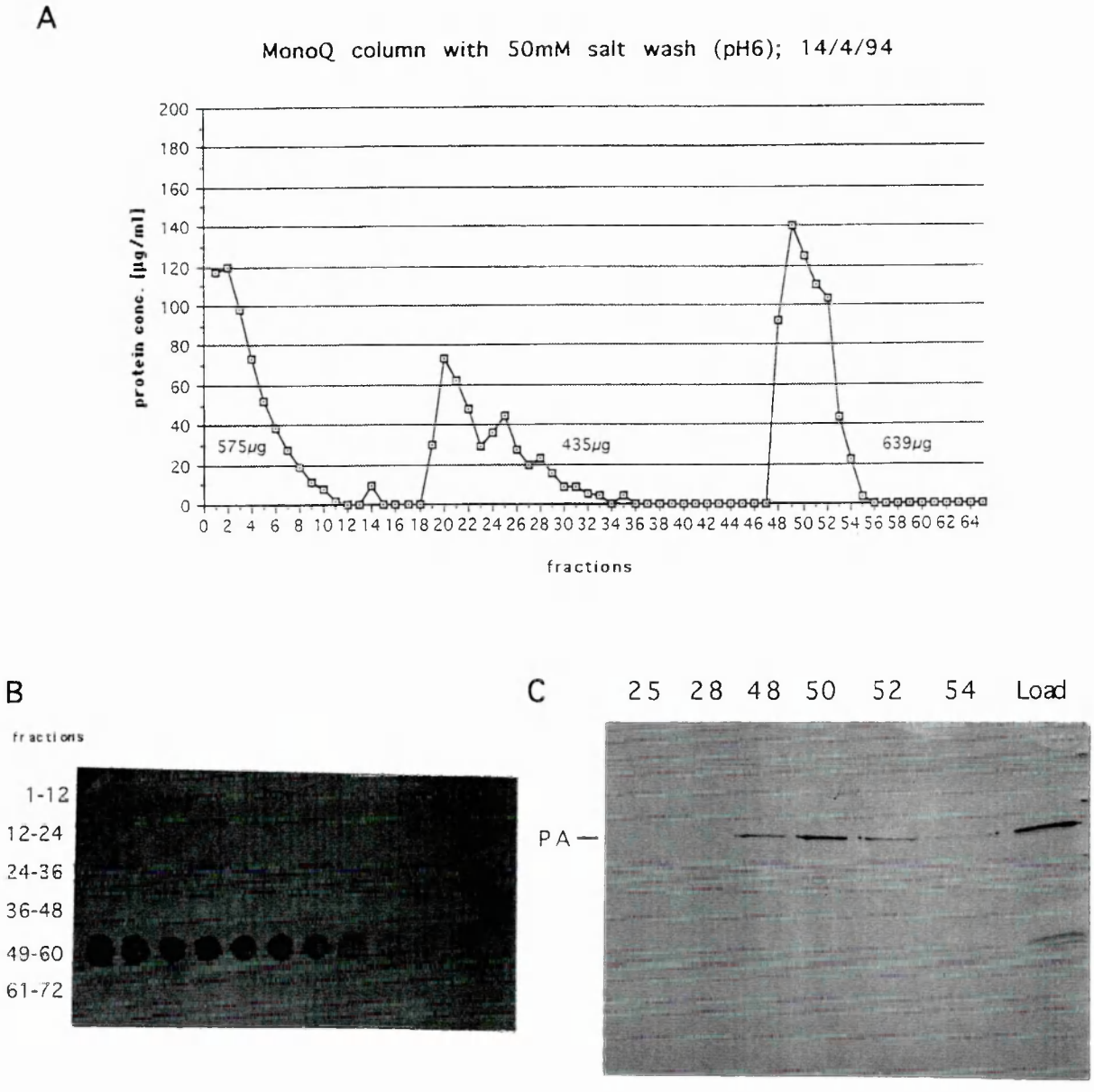


Figure II.12: Binding profile of baculovirus expressed PA protein on a MonoQ column. (A), Protein distribution in the column fractions as determined by Bradford staining. The three major protein peaks correspond to flow-through (575 µg), washing (435 µg) and gradient elution (639 µg). Chromatography was performed in buffer A (50 mM Bis-Tris-HCl), pH 6.0, the column was washed with buffer A plus 50 mM NaCl and bound proteins eluted with a 10 ml 0-1.5 M NaCl gradient. (B), dot-blot of the column fractions shown in (A) and developed with L35E12 anti-PA antibody. PA cross-reactive material was only detected in the eluted fractions. (C), Western blot of column fractions as indicated above each lane. Mainly full-length PA was recovered from the MonoQ column

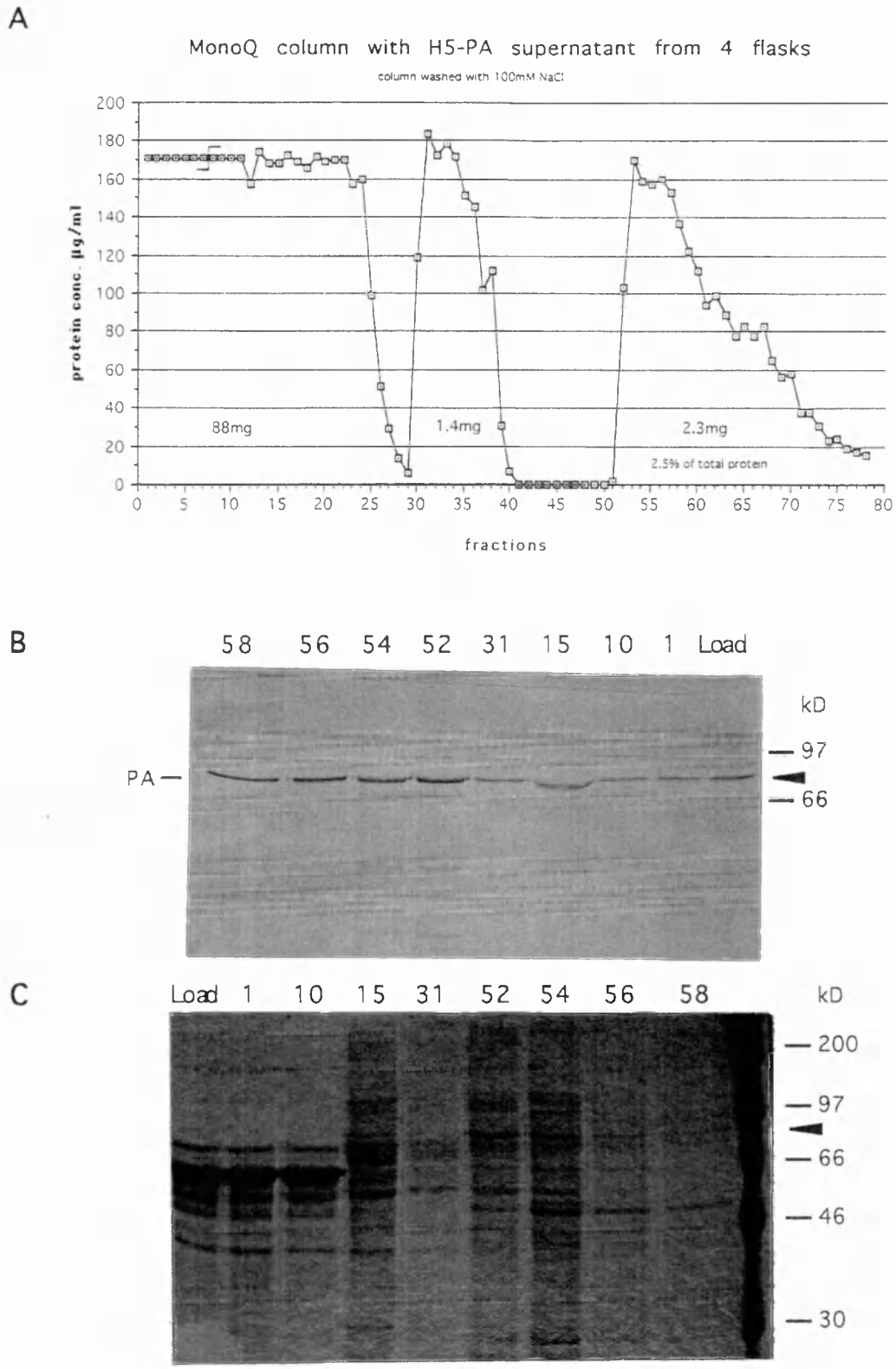


Figure II.13: Purification profile of baculovirus expressed PA from the culture supernatant. (A), protein distribution in the column fractions as determined by Bradford staining; (B), Western blot of column fractions, developed with L35E12 anti-PA antibody; (C), coomassie-stained, 12% SDS-PAGE gel of column fractions as indicated. The arrowhead indicates a band, which corresponds in size to PA and is accumulated in the PA containing peak fractions. Chromatography was performed in buffer A (50 mM Bis-Tris-HCl, pH 6), the column washed with 100 mM NaCl in buffer A and bound proteins eluted with a 20 ml 0-1.5 M NaCl gradient.

II.2.2.4 NTPase activity copurifies with PA on MonoQ and reactive dye columns

The PA containing, semipurified fractions were assayed for ATPase activity as described before for the affinity purified preparations from bacterial expression (II.2.1.3). The MonoQ, Reactive Red and Reactive Green PA peak fractions all contained ATPase activity. Figure II.14 shows that several nucleotide triphosphates (NTPs) could serve as substrate for this activity. α-labeled NTPs were exclusively transformed into labeled NDPs, which demonstrates that all phosphatase activity from the cell extracts was separated from the PA peak fractions in one MonoQ chromatography step. The spot intensities differ for the different nucleotides because they were of different specific activity and equimolar amounts were used in the assay. The NTPase activity in the PA peak fraction utilized ATP and UTP with the highest efficiency, followed by GTP. The incubation of CTP and dCTP with theses protein preparations gave rise to a low concentration of a labeled product of intermediate size between NDP and free phosphate. This product could not be identified.

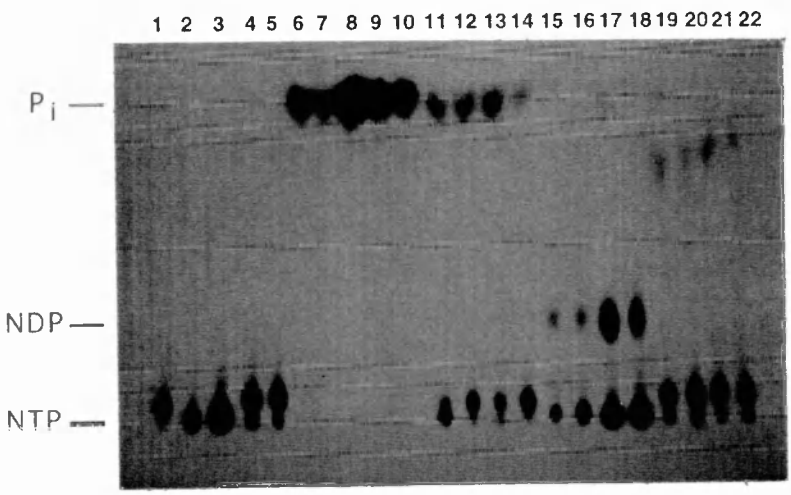


Figure II.14: NTPase activity assay with a PA peak elution fraction from a MonoQ column, corresponding to fraction 50 in figure II.12 and containing an estimated protein content of 0.2 μg/μl. 0.2 and 1 μg protein was incubated with α or γ- labeled NTP in ATPase buffer. 1/20 of the reaction was analyzed by thin-layer chromatography in 1 M LiCl, 0.5 M formic acid. Lanes, **1-5**, incubation of NTPs (γ^{32}P-ATP, α^{32}P-GTP, α^{32}P-UTP, α^{32}P-CTP, α^{32}P-dCTP) in the absence of protein; lanes **6-10**, incubation of the nucleotides in the same order with 0.1 U alkaline phosphatase; lanes **11** and **12**, incubation of 1μg PA peak protein with γ^{32}P-ATP in the presence of nucleotide analogs, 4μM 2'-FdGTP and 2'-FdCTP respectively; lanes **13** and **14**, incubation of 1μg and 0.2 μg PA peak protein with γ^{32}P-ATP; lanes **15** and **16**, incubation of 1μg and 0.2 μg PA peak protein with α^{32}P-GTP; lanes **17** and **18**, incubation of 1μg and 0.2 μg PA peak protein with α^{32}P-UTP; lanes **19** and **20**, incubation of 1μg and 0.2 μg PA peak protein with α^{32}P-CTP; lanes **21** and **22**, incubation of 1μg and 0.2 μg PA peak protein with α^{2}P-dCTP. The positions of marker molecules are indicated on the left.

These results correlate well with the measurements performed on *E.coli* derived PA fractions (see II.2.1.3). It remains therefore possible that PA possesses NTPase activity or associates with an NTPase in both cell types. With appropriate antibodies available for immunopurifications it may be worth trying to determine the nature of this NTPase activity.

II.2.2.5 Conclusion on PA expression in the baculovirus system

The recombinant baculoviruses, that have been produced in this study, express the PA protein at low levels, but in full-length form. Cell disruption can induce very fast and extensive degradation of the PA protein, but the protein appears to be stable after an initial capture step performed with MonoQ anion exchange chromatography. Any analytical work on PA function and activity with recombinant protein from this system will require such a separation of proteases from the PA protein fraction. The purification of PA was attempted from baculovirus infected cells, but the expression levels are too low to obtain pure PA protein by conventional methods or to envisage upscaling of the expression procedure. The conditions worked out for anion exchange chromatography of the PA protein may be important for future purification trials of PA expressed in engineered form and/or in alternative expression systems. Considering the fact that baculovirus infected cells produce mainly full-length PA protein, immunoprecipitation or immunoaffinity purification of PA might be a promising way to obtain reasonably pure protein preparations for the analysis of possible PA activities and for controlled protease treatment to identify independent folding units.

II.2.3 Expression of PA in *Pichia pastoris*

From the above mentioned results it appears that eucaryotic expression systems may contain essential factors required to obtain full-length PA protein. The yields of recombinant protein in baculovirus infected insect cells were low, but significantly higher than in the vaccinia virus system albeit not sufficient for the purification of PA protein. Because exceptionally high amounts of recombinant protein have been reported to be synthesized in *Pichia pastoris* in certain cases, this system was considered as an alternative for PA expression in eucaryotic cells. Recombinant yeast clones were produced by homologous recombination of the yeast genome with a transfer vector, similar to the baculovirus system and following the instructions of the Invitrogen Pichia cloning handbook. The genomes of recombinant yeast clones were analysed by PCR and genomic DNA dot blotting in order to find out the locus and the type of gene integration. Single and multiple insertion clones have been screened for PA expression. The expression level of PA was comparable to the baculovirus system and clearly below 0.5% of total cell protein after 4-5 days of induction. PA was again expressed in mainly full-length form, but the

relative amounts of PA were, as before, too small to achieve efficient enrichment in standard chromatographic steps. Initial trials of protein capturing by Sepharose Q chromatography showed that the protein behaves very similar to baculovirus expressed PA protein, and PA peak fractions from the sepharose Q columns again showed ATPase activity (not shown). However, the yeast system did not improve the yields of recombinant PA as compared to the baculovirus expression system. The amounts are so low that it even remains unclear if a specific affinity purification scheme will be suitable to purify and concentrate PA up to microgram quantitites as would be required for functional studies. The principal advantage of the yeast system is the high cell density to which the cultures grow resulting in high protein concentration per culture volume. Even if the observed PA expression in *Pichia pastoris* was low and with 0.1% similar to the other expression systems, the high total protein production in small volumes (2 g total, soluble protein per liter cell culture) leads to the presence of an estimated 2-5 mg PA protein per liter of cell culture. Considering that upscaling of the cultures is cheaper and easier compared to insect cells, it might eventually be possible to obtain the required amounts of pure protein for functional assays presuming sufficient starting material and a highly specific capturing method like affinity chromatography.

II.2.4 *In vitro* translation of PA protein

The *in vitro* translation system only produces small amounts of recombinant protein, but can be a very important tool in studying specific activities of proteins, especially binding interactions with other proteins or nucleic acids. The available systems are well defined and rapid. The expressed proteins can be specifically labelled by incorporation of radioactive amino acids and background labelling of other proteins is very low. Mutations in the recombinant protein sequences can be produced relatively fast. The expression plasmids are usually stable during propagation in *E.coli*, even if they contain coding sequences of toxic proteins, because viral polymerases are required for protein expression. These are all advantages, that made it worthwile expressing PA in this system in order to study potential RNA binding activity of this protein. The PA protein sequence does not contain any of the described RNA binding motifs (Burd and Dreyfuss, 1994). Nevertheless, it is part of the influenza virus polymerase complex and could be UV-crosslinked to nascent RNA transcripts (Braam et al., 1983). Also, a presumed function in translational control might involve RNA binding. A significant amount of nucleic acid bound to the amylose resins during the purification of MBP-PA fusion proteins as judged by UV absorption spectroscopy. These nucleic acid moieties could be partially removed from PA by washing the columns with elevated NaCl concentrations. All these observations suggest that PA might be a RNA binding protein, and if so, it may contain a novel RNA binding structure. The expression of PA in wheat germ extracts was tested for its use to study PA protein function *in vitro*. To this end, full-length PA and two PA fragments from

the termini of the protein were cloned into T7 polymerase directed expression vectors. An optimal Kozak sequence for efficient translation initiation was added immediately upstream of the start codon during the cloning procedure (Cavener and Stuart, 1991). The coding sequences were transcribed and translated in the combined TNT™ system of Promega and analyzed for RNA binding activity by measuring the binding of ^{35}S-labelled PA to biotinylated RNA molecules according to a procedure originally described by Scherly et al., 1989. RNA-bound protein is then detected after precipitation of the RNA substrates with streptavidin coated magnetic beads.

The RNAs, that have been used in these binding experiments, have either been chemically synthesized and biotinylated by the oligonucleotide service group at EMBL, Heidelberg, or they have been *in vitro* transcribed with the concomitant incorporation of biotinylated UTP (table II.2, page 74). The magnetic Dynabeads (DYNAL) were blocked by a preincubation in KTN-buffer containing 10 mg/ml *E.coli* tRNA and 1.5 mg/ml BSA to avoid nonspecific interactions of the PA protein with the streptavidin beads. For RNA binding, in vitro translated PA protein was incubated with 2 μM biotinylated RNA in the presence of 300 μM polyA RNA for 10-30 min at RT. The binding reaction was then diluted 20-fold with KTN-buffer, 0.15 μg of magnetic beads were added and left 30 min at RT for RNA capturing. The beads were washed three times with 20-fold excess KTN-buffer, the bound material was eluted in SDS-loading buffer and analyzed by 12 % SDS-PAGE and autoradiography. In this assay, the PA protein showed selective binding to some, but not all RNAs tested. Figure II.15 shows the coprecipitation of *in vitro* translated PA with different, biotinylated RNA molecules. The blocking of the beads with unspecific RNA and BSA resulted in very low background binding of PA to the beads in the absence of RNA or in the presence of non-biotinylated RNA (figure II.15A, lanes 3, 5, 10, 13 from left to right). PA was found to precipitate with a RNA oligonucleotide containing the conserved 5' end sequence of influenza virus vRNA (con5), but not with the conserved 3' end sequence (PAV3 and con3). PAV3 corresponds to the 3' end of influenza A virus segments 1,2,3,5 and 7, con3 to the 3' ends of segments 4, 6 and 8 (see figure I.14, page 31). In addition, PA apparently associated specifically with the RNA oligonucleotide "oNS-mRNA" (figure II.15A, lane 7). oNS-mRNA represents the 5' end of influenza virus specific mRNA coding for NS1 protein. The coprecipitation of PA with oNS-mRNA was independent of a cap structure and presumably mediated by the influenza virus-specific sequence, because pucT7bam RNA did not precipitate PA in spite of containing the primer sequence present in oNS-mRNA at the 5' end. Also, biotinylated U6snRNA did not precipitate PA from *in vitro* translation extracts.

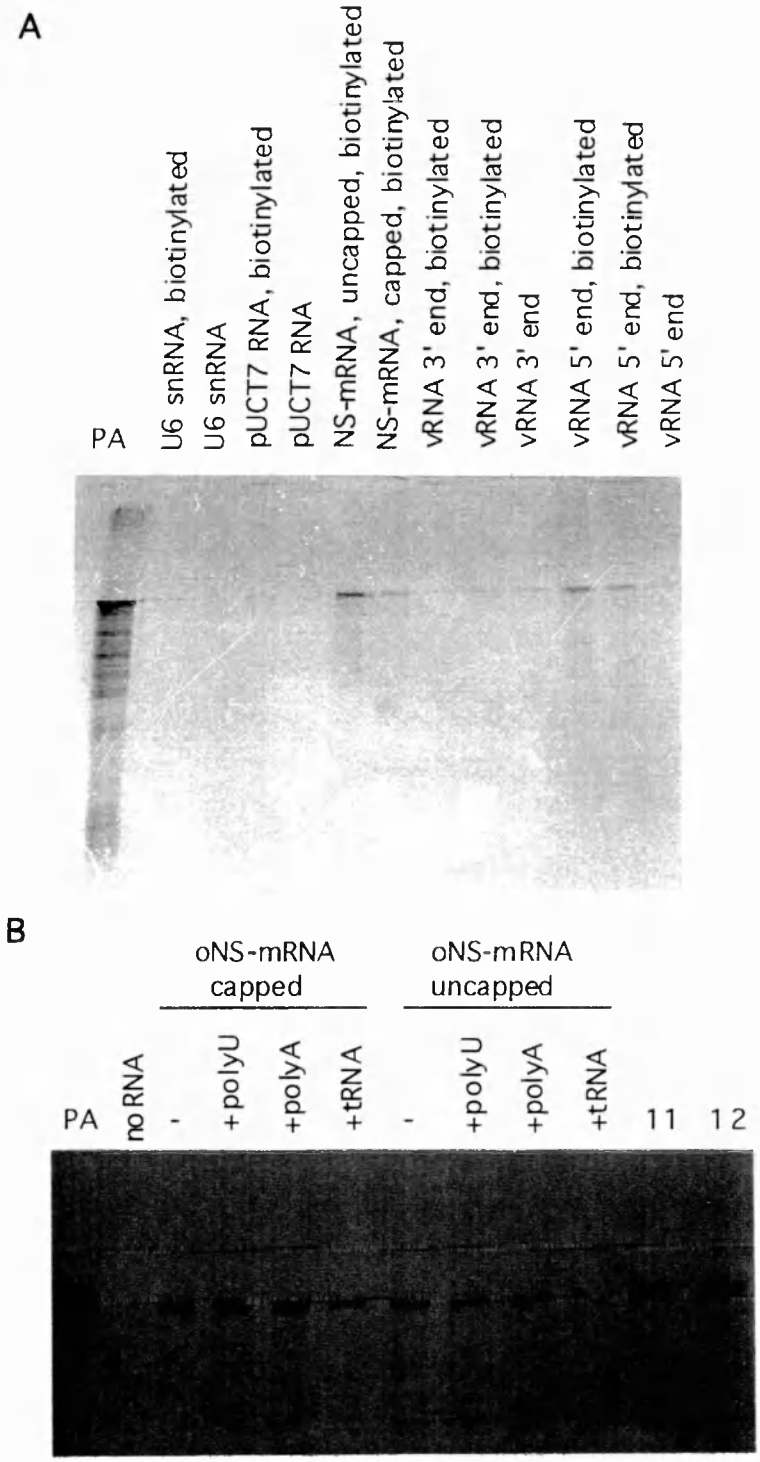


Figure II.15: RNA binding of *in vitro* translated PA protein. 12% SDS-PAGE autoradiograms showing coprecipitation of radioactively labeled PA protein with biotinylated RNAs. (A), specific binding of PA to vRNA 5' end (con5) and influenza virus mRNA. The RNAs used in the binding assay are indicated above the corresponding lanes. "PA" is an acetone precipitation of 1/10 of the protein in the binding reaction. (B), Cap independence and effect of non-specific competitors on PA binding to NS-mRNA. "-" indicates the absence of competitor in the reaction. Lanes 11 and 12 are binding reactions with oNS-mRNA, capped and uncapped, that has been polyadenylated *in vitro* with E.coli poly(A)-polymerase (Pharmacia).

Figure II.15B illustrates that the RNA dependent precipitation of PA was not inhibited by up to 625 fold molar excess of polyA, polyU or tRNA. Similar results were obtained with capped or uncapped mRNA analogs. Binding was not changed by the presence or absence of a polyA tail added to the mRNA analogs *in vitro*, which is consistent with the absence of competition by polyA RNA. These results suggest that PA might specifically bind to influenza virus specific sequences from the 5' ends of vRNA and cRNA, but not to the 3' end of vRNA. The differential binding of full-length PA to the RNAs implies specificity of the binding activity. The point mutant PAL219F showed similar binding characteristics to wild-type PA, but preliminary results with the PA-NT and PA-CT deletion mutants suggests very strongly reduced binding in this assay.

Taken together, the expression of PA in wheat germ extracts provided PA protein preparations, that may be used to study RNA binding *in vitro*. The results so far are consistent with the idea of a function of the PA protein in isolated form. However, others have failed to detect any specific RNA binding activity of isolated influenza virus polymerase subunits expressed in the vaccinia virus system. (Tiley et al., 1994). The specificity of this assay will have to be scrutinized and RNA binding be shown with yet another method, e.g. gel shift or UV-crosslinking. If RNA binding activity of PA can be established with this system, PA point and deletion mutants will possibly help to identify the RNA binding site on the PA primary sequence. On the other hand, a mutational analysis of the RNA molecules will help to determine the character and specificity of the RNA binding site.

RNA	sequence	length [nt]	origin	PA binding
con 5	5' bio-AGU AGA AAC AAG GGT G 3'	16	chem. syn.	+[a]
con 3	5' bio-CAC CCU GCU UUU GCU 3'	15	chem. syn.	-
oNS-mRNA	5' GGG CGA UCA AGA GCA AAA GCA GGG UGG CAA AT-bio 3'	32	chem. syn. capping reaction[b]	+
PAV 3	5' bio-GUA CCU GCU UUC GCU 3'	16	chem. syn.	-
U6 snRNA	5' GGGUGCUUGCUUCGGCAGCAC AUAUACUAAAAUUGGAACGAUAC AGAGAAGAUUAGCAUGGCCCCUG CGCAAGGAUGACACGCAAAUUCG UGAAGCGUUCCAUAUUUUU 3'	109	T7 transcription from *Dra* I cut pUC9U6 plasmid[c] (in the presence of biotinylated UTP)	-
pucT7bam	5' GG GCG AUC AAG CUA UGC AUC UCG AGG AAU UCG 3'	32	T7 transcription from *Bam* HI cut pUCT7 plasmid[d] (in the presence of biotinylated UTP)	-

Table II.2: Characteristics of the RNAs, that have been used in PA binding studies.

[a] PA binding determined by precipitation with streptavidin beads.

[b] In vitro capping of RNAs was performed with vaccinia virus guanylyl-transferase as described in chapter IV (Mat/Met).

[c] Plasmid pUC9U6 was obtained from Dr. Iain Mattaj and contains the coding sequence for *Xenopus tropicalis* U6snRNA downstream of a T7 promoter

[d] Plasmid pUCT7 was constructed by ligating the annealed DNA oligonucleotides T7PGEMTOP and T7PGEMBOTTOM into *Hin*dIII/*Bam*HI cut pUC18 plasmid.

II.3. DISCUSSION

Very little is known about the structure-function relationships of RNA dependent RNA polymerases, which are the central enzymes in the life cycles of RNA viruses. On the one hand, they contain highly conserved domains present in most RNA polymerases and they seem to follow basic rules of templated RNA synthesis, as established for cellular RNA polymerases. On the other hand, they show virtually no sequence similarities in other protein domains apart from the polymerase module, and they usually perform very unique, virus-specific activities required for genome transcription and replication. These

additional activities potentially constitute very powerful targets for antiviral therapies. The viral RNA-dependent RNA polymerases are large protein complexes (> 200 kD), sometimes including a variety of cellular proteins (especially in positive-strand RNA virus groups). The rules that govern the assembly of functional polymerase complexes are not well understood, partly because of considerable difficulties in expressing single subunits and the reconstitution of specific activities *in vitro*.

The goal of the present study was to assess a number of protein expression systems to try to find a way to produce functional influenza virus polymerase subunits for structural studies. In addition, we attempted to detect a specific activity of the PA subunit to help understand the function of this protein in the virus life cycle. In most cases the production of full-length, recombinant PA protein was below 0.5% of total cellular protein. A number of reasons could theoretically account for this observation and it has to be discussed, if any promising strategies to increase protein yield became apparent from the experiments. As a first obstacle, plasmid instability was observed during overnight cultures of *E.coli* cells carrying recombinant plasmids encoding PA proteins. The comparative screening of colony formation on agar plates with different additives suggested that an overnight culture grown until OD_{600} = 2 could lose the PA expression plasmid from more than 66% of the cells. A similar segregation process has been described for the expression of toxic gene products in bacterial cells (Novagen pET system manual). The most likely reason for this phenomenon is the production and secretion of large amounts of β-lactamase in *E.coli* cell cultures carrying the recombinant plasmids. Accordingly, ampicillin selection is lost relatively early in cell culture. Colonies, that have lost the plasmid, will start to overgrow the cell cultures as soon as sufficient β-lactamase has accumulated, especially in cases, where the retention of the plasmid is unfavourable for the cells. To suppress this problem, cell cultures for plasmid production, before protein induction, were not allowed to grow further than OD_{600} = 0.5. The fraction of plasmid containing cells was monitored during the cultures to minimize unspecific overgrowth. However, it was difficult to avoid plasmid loss during protein induction itself. After the addition of IPTG at OD_{600} < 0.5, plasmid containing cells virtually ceased to grow and started to lyse 2 hours later, as apparent by a decrease of the optic density (OD_{600}). Protein induction at higher optic densities had either similar effects or increases of OD_{600} levels were measured, which had resulted from overgrowth by plasmid-less cells, as could be seen in plate screening assays. These observations suggest that the expression of PA is deleterious for the bacteria. This apparent toxicity of the PA gene product might possibly indicate that PA possesses an activity in isolated form and that this activity is deleterious for bacterial cells.

Most of the recombinant PA protein was produced in an insoluble form from the pRSET and pinpoint systems, both of which add short extra-sequences to the N-terminus of PA. Similar results have been obtained with different *E.coli* expression systems in

other laboratories (Akkina et al., 1987; Barcena et al., 1994; Jones et al., 1986; Kobayashi et al., 1992). Consistent with previously published results, the yields of full-length proteins were very low, whereas certain protein fragments could be produced to much higher levels. We did not succeed in refolding PA or smaller PA fragments into functional structures after the purification of PA from bacterial inclusion bodies. By fusing the maltose binding protein (MBP) to the N-terminus of PA, it could for the first time be produced in soluble form in *E.coli*. However, PA was largely fragmented into polypeptides of all sizes and the yield of full-length PA itself was again very low. The addition of a large variety of protease inhibitors, single and in combinations, did not significantly reduce the extent of protein fragmentation. The fragments were also detected in whole cell extracts prepared from intact cells solubilized in SDS-loading buffer. We conclude that to a large extent this fragmentation already occurs *in vivo* and the contribution of protein degradation during cell lysis was not easy to estimate on this background. High level degradation *in vivo* might be due to problems during protein folding, because proteins with abnormal conformations are known to be funneled into protein degradation pathways in *E.coli* (Goldberg, 1992; Gottesmann and Maurizi, 1992). Coexpression of molecular chaperones (e.g. DnaK, GroEL, GroES) could be used to answer this question. However, these chaperones have been reported to stimulate proteolysis in *E.coli* with polypeptides, that fail to develop a stable conformation. Therefore, this strategy does not generally lead to improved protein expression (Kandror et al., 1994). A second reason for intracellular fragmentation of PA might be inefficient processivity during transcription or mRNA instability. In *E.coli*, mRNA turnover is regulated by the coordinated action of endo- and exonucleases. The best studied pathway starts with the initial cleavage of mRNAs near the 3' end by endonuclease RNase E (Hajnsdorf et al., 1994). The stability and half-life of the PA mRNAs in the respective expression systems has not been looked at in this study, but several, putative RNAse E cleavage sites, that are common to extremely unstable mRNAs, are also present in the PA gene sequence. An RNAse E-like activity has also been described to be important in the control of mRNA decay in eucaryotes. Repeated AUUUA motifs and AU-rich regions are common in human oncogene and growth factor mRNAs and significantly reduce the half-life of the mRNAs (Wennborg et al., 1995). Figure II.16 shows the distribution of possible RNase E recognition sites on the PA gene sequence. The conservation of these signals between bacteria and eucaryotes may partly explain the low expression levels of PA in both procaryotic and eucaryotic protein expression systems.

PA			RNase E consensus cleavage site
406	**AAUAAAAUUAAAT**	418	
680	**AAAAUUUUAGA**	690	$\frac{A}{G}$ A U U $\frac{A}{U}$
862	**UUAAAAUUAA**	872	
996	**AAAUUAU**	1002	
1442	**AAUUAAUU**	1449	
1713	**AAAAAUUAAAAUGAAAU**	1729	
1935	**UAUUUAA**	1940	
2081	**UAAUUAAUGAU**	2091	

Figure II.16: Motifs in the PA gene sequence (IA/PR/8/34), that resemble RNase E cleavage sites implied in the regulation of mRNA stability in both procaryotic and eucaryotic cells (Ehretsmann et al., 1992; Mackie and Genereaux, 1993; Wennborg et al., 1995). Very low half-life mRNAs are characterized by clusters of the RNAse E recognition motif.

As mentioned above (II.2.1.2), the presence of large numbers of rare codons in eucaryotic genes has also been suggested to reduce the levels of protein production in bacteria and can lead to protein fragmentation. Table II.3 shows the codon usage of PA compared to highly expressed genes in *E.coli* and in *Saccharomyces cerevisiae*. The most striking deviations are found in the codon redundancy families for arginine and glycine. The codons that are used in the PA gene for these amino acids are only rarely used in *E.coli* and *S.cerevisiae* with the exception of the AGA codon, that is common in yeast genes. In addition, the codons that are preferentially used in the PA gene for leucine, threonine and asparagine are rare codons in the microbial organisms. A helper plasmid expressing the tRNA for the rare arginine codons has been reported to improve the expression of various genes of plants, plant viruses and *Agrobacterium* in different laboratories (Schenk et al., 1995). The AGA/AGG contents of these genes ranged between 1.8% and 7.8%, which is comparable to PA with 4.4%. Indeed, the coexpression of the $tRNA_{arg4}$ also increased the level of PA production in *E.coli*, but did not significantly reduce the level of PA fragmentation. For the future it might therefore be worth trying to coexpress all rare tRNAs together with PA in bacterial cells.

C	AA	PA	E.c.	S.c.	C	AA	PA	E.c.	S.c.	C	AA	PA	E.c.	S.c.	C	AA	PA	E.c.	S.c.
TTT	F	0.44	0.26	0.11	TCT	S	0.19	0.36	0.57	TAT	Y	0.65	0.26	0.06	TGT	C	0.4	0.36	0.93
TTC	F	0.56	0.74	0.89	TCC	S	0.17	0.34	0.38	TAC	Y	0.35	0.74	0.94	TGC	C	0.6	0.64	0.07
TTA	L	0.11	0.03	0.11	TCA	S	0.21	0.02	0.01	TAA	Z				TGC	Z			
TTG	L	0.17	0.04	0.85	TCG	S	0.08	0.04	0.00	TAG	Z				TGG	W	1.00	1.00	1.00
CTT	L	0.34	0.06	0.01	CCT	P	0.13	0.10	0.06	CAT	H	0.15	0.20	0.19	CGT	R	0.05	0.67	0.09
CTC	L	0.09	0.08	0.00	CCC	P	0.19	0.01	0.01	CAC	H	0.85	0.80	0.81	CGC	R	0.05	0.30	0.00
CTA	L	0.12	0.01	0.03	CCA	P	0.48	0.12	0.93	CAA	Q	0.58	0.18	0.98	CGA	R	0.08	0.01	0.00
CTG	L	0.17	0.79	0.00	CCG	P	0.19	0.77	0.00	CAG	Q	0.42	0.82	0.02	CGG	R	0.03	0.00	0.00
ATT	I	0.48	0.20	0.50	ACT	T	0.24	0.36	0.48	AAT	N	0.65	0.08	0.06	AGT	S	0.19	0.04	0.02
ATC	I	0.25	0.79	0.50	ACC	T	0.19	0.48	0.51	AAC	N	0.35	0.92	0.94	AGC	S	0.17	0.19	0.02
ATA	I	0.27	0.01	0.00	ACA	T	0.57	0.05	0.01	AAA	K	0.58	0.71	0.10	AGA	R	0.45	0.01	0.91
ATG	M	1.00	1.00	1.00	ACG	T	0.00	0.10	0.00	AAG	K	0.42	0.29	0.90	AGG	R	0.35	0.00	0.00
GTT	V	0.23	0.44	0.55	GCT	A	0.21	0.37	0.76	GAT	D	0.64	0.36	0.39	GGT	G	0.11	0.54	0.98
GTC	V	0.10	0.09	0.44	GCC	A	0.29	0.10	0.23	GAC	D	0.36	0.64	0.61	GGC	G	0.14	0.41	0.02
GTA	V	0.37	0.27	0.00	GCA	A	0.49	0.25	0.01	GAA	E	0.55	0.76	0.98	GGA	G	0.51	0.01	0.00
GTG	V	0.30	0.19	0.01	GCG	A	0.01	0.28	0.00	GAG	E	0.45	0.24	0.02	GGG	G	0.26	0.03	0.00

Table II.3: Codon usage of the influenza A/PR/8/34 PA virus gene (PA) as the fraction of each codon in its redundancy family, compared to strongly expressed genes in *E.coli* (E.c.) (Grosjean and Fiers, 1982) and *Saccharomyces cerevisiae* (S.c.) (Sharp and Cowe, 1991). Boxes indicate a significant deviation in the codon preference of the PA gene. "C" stands for codon, "AA" for amino acid. Amino acids are denoted in the standard one-letter code.

PA fragmentation *in vivo* was not a problem in the eucaryotic expression systems. In both insect and yeast cells, PA was mainly produced as a full-length protein and smaller bands detected on Western blots were generally derived from proteolysis after cell lysis, which could be minimized by appropriate harvesting procedures. However, the amount of PA protein was very low and comparable to the *E.coli* systems, i.e. PA protein accumulated to about 0.1% of total cellular protein. This means that roughly 0.2-1 mg PA was produced per liter of insect cell culture and 2 mg per liter of yeast culture. The low level expression in all systems suggests that the signals preventing the accumulation of PA are conserved between bacteria and eucaryotes.

The PA protein may be toxic in both cell types. The plasmid instability and decrease in OD_{600} after induction in *E.coli* are signs for toxicity. The yeast cell cultures expressing PA showed reproducibly and significantly reduced growth compared to control cells expressing a recombinant tRNA synthetase. Possible toxicity in insect cells is hard to determine, because during the late phase of infection the cells start to be lysed by the replicating baculovirus. The PA expressing baculoviruses did not grow to very high titers

and considerably less recombinant virus clones were obtained with PA compared to PB1 and PB2, which both may be signs for PA toxicity. Similar observations have been reported from other laboratories as well. Trials to produce stable cell lines expressing PA from A2G mouse kidney tumor cells were unsuccessful, because transfected cells stopped growing after a few days (Stranden et al., 1993). Nieto et al., (1992) reported that nuclear location of PA correlated with chromatin condensation and aberrant nuclear morphology. One reason for this toxicity may be the induction of a generalized degradation of proteins induced by the expression of PA and shown to occur in different cell lines (Sanz-Ezquerro et al., 1995). The mechanism and the physiological role of this induction of proteolysis are not known. Since the N-terminus of PA appears to be required for this activity (Sanz-Ezquerro et al., 1996), in order to reduce toxicity and increase expression efficiency it might be worthwile to try and express N-terminally truncated PA proteins and see if they can be accumulated to higher concentrations and if they retain certain activities like interaction with other polymerase subunits and RNA binding (see below). However, it is not known if similar reasons can account for toxicity in bacterial cells.

Attempts by other groups to produce recombinant influenza virus polymerase subunits in insect cells were similarly hampered by low expression levels and to my knowledge no one has succeeded so far in obtaining pure and functional single subunits (Kobayashi et al., 1992; Shi et al., 1995; St. Angelo et al., 1987). The complexity of possible reasons for low level expression of PA makes it unlikely that simple strategies exist to significantly increase the yield of recombinant PA protein. Only highly specific affinity purification approaches might provide a means to obtain recombinant protein for functional studies.

The PA preparations from different systems showed similar features in ATP binding and ATPase activities. A certain specificity of these activities was implied by the absence of activity in protein preparations from control cell extracts that had expressed either no protein, an unspecific protein or a N-terminal fragment of PA. In addition, the binding was specific for ATP, as none of the other nucleotide-triphosphates was bound or could compete for ATP binding. Also, no phosphatases copurified with PA. However, in *E.coli* the ATP-binding activity was correlated to a specific breakdown product of the full-length MBP-PA fusion protein, whereas in the eucaryotic systems the degree of purification was not sufficient to completely exclude the copurification of an unspecific ATPase with PA. To determine the nature of the ATPase activity, different anti-PA antibodies or antisera could be screened for specific inhibition or copreciptation of protein and activity. Alternatively, point mutations in the presumed ATP binding motif of PA might help to correlate ATPase activity to the PA protein.

PA expression in wheat germ extracts was evaluated as a means to study RNA binding activity. The efficiency of PA coprecipitation with specific RNAs was low, but comparable to previous work by Scherly et al. (1989). The present study suggests that *in vitro* translation of the influenza virus polymerase subunits in different combinations and the production of mutant proteins in this system might be an alternative to the recombinant vaccinia virus system and constitute a promising way to obtain more information about the RNA binding determinants of the influenza virus polymerase. As an unexpected first result we obtained evidence that the PA protein alone can specifically bind to the conserved sequence at the 5' end of influenza virus vRNA and mRNA. This dual binding specificity may be conferred by the very similar sequence of these RNA ends and may explain how one protein could have a function in the genome binding of the polymerase complex and outside the complex in virus specific gene expression. In the future it could be tried to determine the physiologic role of this activity and establish the crucial parameters on the RNA and polypeptide sequences. A similar approach has resulted in the functional characterization of other eucaryotic proteins, that were difficult to express in standard protein expression systems. The determination of the minimal RNA binding domain of the spliceosomal U1A protein provided the basis for efficient expression and crystallization of the first RNA-binding-domain in the presence of its cognate RNA (Oubridge et al., 1994; Scherly et al., 1990; Scherly et al., 1989). A previously unknown RNA-binding activity was shown with this technique for the FMR1 protein, and the loss of the activity seems to be correlated with the development of fragile X syndrome in humans (Ashley et al., 1993). Finally, the interaction of the ribosomal L5 protein with 5S RNA was studied in the *in vitro* translation/biotinylated RNA system to characterize this complex, which is important in nucleolar ribosome assembly (Michael and Dreyfuss, 1996).

Taken together, all expression systems that have been tested in the present study are unlikely to produce sufficient full-length PA protein for crystallographic studies. The most promising strategy for the future determination of PA structure and function will be to determine functional domains in this polypeptide that are easier to express at high levels. The in vitro translation approach may help to define such domains. An alternative approach might be to produce sufficient PA in insect or yeast cells for affinity purification, proteolytic studies and microsequencing of stable intermediates as a basis to identify the regions to be expressed in isolated form. Preliminary functional characterization of PA from different expression systems have suggested possible novel activities for this protein. At present, many further experiments are still needed to unequivocally establish and characterize these activities. However, an activity assay looking at complex formation, RNA binding or ATPase activity is crucial for the development of PA fragment expression.

Chapter III

In vitro transcription activity of influenza virus RNPs

III.1 Introduction

The influenza virus possesses a unique, RNA-dependent RNA polymerase, which is a multifunctional protein complex consisting of the three subunits PB1, PB2 and PA. As outlined above (I.4.2), two different forms of the polymerase are presumably involved in transcription and replication events. The polymerase found in virus particles shows only transcription activity, which is characterized by the use of 5' capped (m7GpppNm-containing) oligoribonucleotides as primers, and termination of RNA synthesis 17-22 nucleotides before the 5' ends of the vRNA templates. Polymerases capable of replication *in vitro*, i.e. primerless initiation of RNA synthesis and production of full-length copies of the template RNAs, have only been found in influenza virus infected mammalian cells at certain time-points after infection (Krug et al., 1989). Whereas the polymerase complex presumably carries all enzymatic activities required for transcription and replication, complete RNPs are the actual functional units. The packaging of the template RNA into distinct structures and the presentation of the bases by nucleoprotein (NP) seem to be required for efficient elongation during RNA synthesis. The RNPs, that are present in virus particles, show several activities *in vitro*, which can potentially be used as markers for the screening of potential antiviral compounds. They possess endonuclease activity, which means that they can characteristically cleave a variety of capped RNAs at distinct positions (Plotch et al., 1981). Transcription is initiated by the addition of a GTP or a CTP to a capped primer RNA. The transcripts are then elongated by the incorporation of further nucleotides. Finally, transcription termination and polyadenylation occur at an oligoU stretch near the 5' end of the vRNA (Hay et al., 1977a; Hay et al., 1977b; Plotch and Krug, 1977; Robertson et al., 1981). *In vitro*, the RNPs possess a low level transcription activity even in the absence of primer RNA, which can be stimulated 10-100 fold by the addition of the dinucleotides ApG or GpG (McGeoch and Kitron, 1975; Plotch and Krug, 1977). The dinucleotides are incorporated into the transcripts at transcription initiation and the RNA products are polyadenylated and can be translated in cell-free systems (Bouloy et al., 1978; Plotch and Krug, 1978). Complete RNPs, purified from egg-grown influenza A virus, have been used in this study to determine the mechanism of inhibition by 2'-fluororibonucleotides, which show considerable antiviral activity in cell culture and in the mouse pneumonia model (Tisdale et al., 1993).

III.2 Principle of influenza virus transcription assay systems

Since the discovery of polymerase activity in influenza viruses, a number of different *in vitro* transcription assay systems have been developed to study polymerase activities, promoter structures and the mechanism of inhibition by anti-polymerase

compounds. In most cases, complete viral RNPs from detergent disrupted influenza virus are used as the polymerase source. The virus can be grown in fertilized chicken eggs or in mammalian cell cultures and be concentrated by sucrose gradient centrifugation. Lysis of the viral membrane is achieved in the presence of 0.2-1% Triton N101 or Triton X100 and 1% Lysolecithin, and RNPs can be further purified by glycerol gradient centrifugation (Honda et al., 1988; Plotch and Krug, 1977; Rochovansky, 1976). Both disrupted viruses and purified RNPs show similar transcription activity *in vitro*, which can be measured by the incorporation of radioactively labeled nucleotide-monophosphates (NMPs) into trichloroacetic acid (TCA) insoluble material. Complete viral mRNAs are only produced in the presence of specific primers, either dinucleotides or capped and methylated mRNAs (Bouloy et al., 1979; McGeoch and Kitron, 1975; Plotch et al., 1979; Plotch et al., 1981; Plotch and Krug, 1977). The dinucleotide ApG primes transcription presumably by base-pairing with the 3' terminal bases of the vRNA template and is incorporated into transcripts (Honda et al., 1986). Labeled products of *in vitro* transcription can be resolved on denaturing urea-polyacrylamide gels. Relatively sharp bands are obtained after the removal of the polyA tails by RNAse H digestion in the presence of polyU RNA. These bands, after ApG primed transcription, migrate slightly faster than the vRNA molecules, because they are 17-22 nucleotides smaller than their templates (Plotch et al., 1979; Plotch and Krug, 1978). With labeled mRNAs as primers of influenza virus transcription, the endonuclease reaction and transcription initiation can be studied as cleavage and elongation of the precursor molecule. The reaction products are again visualized by denaturing gel electrophoresis (Plotch et al., 1981). In this study, transcription initiation and elongation were studied using *in vitro* transcribed and capped RNA as primer substrate together with labeled nucleotides and purified RNPs.

III.3 RNA substrates for influenza virus transcription

A large number of different RNAs can serve as substrates for the endonuclease activity of the influenza virus polymerase complex. Several laboratories have sequenced the 5' ends of viral mRNAs produced in infected cells, but the same host cell derived primer sequence has never been observed twice in these clones, suggesting that there is considerable redundancy in the mRNA substrate choice (Lamb and Choppin, 1983). Naturally occuring primer sequences imply a preference for the sequence PyGCA upstream of the cleavage site, but cleavage after G residues is also common. Studies with capped nucleotide homopolymers led to the suggestion that the polymerase might also recognize the nucleotide downstream adjacent to the cleavage site (Kawakami et al., 1983). It appears that the polymerase preferentially uses primers, that are able to undergo base-pairing interaction with the 3'-terminal uridine of the viral RNA template, although base-pairing between primer and template are not required *in vitro* (Beaton and Krug, 1981; Lamb et al., 1981; Hagen et al., 1995; Krug et al., 1980). Consistent with the heterogeneity observed *in vivo*, a number of artificial RNAs can serve as specific

endonuclease substrates *in vitro*, e.g. globin mRNA, reovirus mRNA, several *in vitro* capped plant virus RNAs, capped polyA and polyU and other synthetic RNAs (Bouloy et al., 1979; Bouloy et al., 1980; Plotch et al., 1979; Plotch et al., 1981; Kawakami et al., 1983; Chung et al., 1994). The presence of a methylated cap structure is crucial for RNA molecules to serve as substrates for the endonuclease. Uncapped RNAs or RNAs carrying a 5'-terminal GpppG group are inactive as substrates. RNAs with cap 0 structures (m7G(5')ppp(5')X) are usually active, but priming activity is stimulated 2-30 fold in the presence of additional 2'-O-methyl groups in the cap of mRNAs, so-called cap I structures (m7G(5')ppp(5')Xm) (Bouloy et al., 1980). The distance between cap structure and cleavage site always ranges between 9 and 15 nucleotides. Efficient endonuclease substrates are presumed to possess an optimal sequence surrounding this position, but the mechanism behind the observed cleavage site selectivity is not yet understood. There is a preference for cleavage after purines, but it is unpredictable, which purine will be preferentially used on a given RNA with several purines between position 9-15. Morevover, cleavage after U residues has also been repeatedly observed albeit at lower efficiency (Hagen et al., 1995). In the present study, commercially available rabbit globin mRNA (GIBCO), two synthetic RNAs, the pGEM transcript and the pUCT7GEM RNA, as well as the dinucleotide ApG were used as primers for influenza virus specific *in vitro* transcription reactions. pGEM RNA was produced by SP6 polymerase directed, *in vitro* run-off transcription reactions from *Sma* I digested pGEM7Zf+ plasmid DNA (Promega), pUCT7GEM RNA by T7 polymerase directed transcription from pUCT7GEM plasmid (sequence, see table II.3, page 78). A cap 0 structure could be added with approximately 50% efficiency by including a cap analog (m7GpppGTP) in the *in vitro* transcription reaction. Alternatively, an *in vitro* capping reaction was performed on gel-purified RNAs using vaccinia virus guanylyltransferase, which has very recently been made commercially available (GIBCO). Unlike other mRNA substrates, which are usually cleaved heterogeneously by the influenza virus polymerase, the capped pGEM RNA is specifically cut at a single position, after a G residue, 11 nucleotides downstream of the cap. The resulting endonuclease product is called G11 (Chung et al., 1994). This unique quality of the pGEM RNA facilitates the identification of transcription products on analytical gels. Figure III.1 schematically depicts the presumed base-pairing interactions between template and nascent RNA during transcription initiation and the products of transcription initiation reactions in the presence of CTP and GTP.

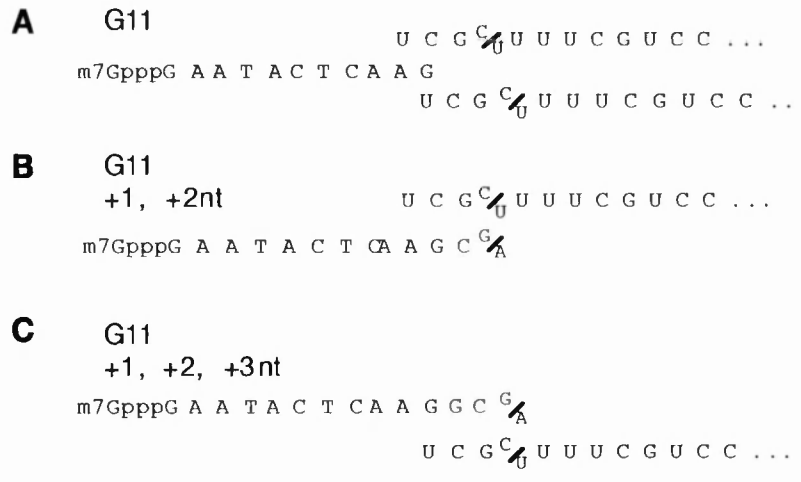


Figure III.1: Products of RNP directed transcription initiation in the presence of CTP and GTP and pGEM RNA as primer. (A) shows the proposed base-pairing interaction between the endonuclease product and the genomic RNA template. Transcription can be initiated with CTP (B) or GTP (C) to form a 12 nt product. In the presence of both GTP and CTP the G11 primer can be elongated with 1, 2 or 3 nucleotides (see figure III.7).

III.4 Influenza virus *in vitro* transcription

III.4.1 Characterization of general transcription conditions

Trichloroacetic acid (TCA) precipitation assays were performed to measure the relative transcription activity of purified RNPs from different influenza A/PR/8/34 virus preparations, that were obtained from Pasteur-Mérieux, Lyon. These assays measure the incorporation of radioactively labeled NMP into acid insoluble material. We found up to 100-fold differences in transcription activity between different virus batches, but we have not yet an explanation for these differences. The K_m values for ribonucleotides, as determined by double reciprocal plots, were similar for the high activity and low activity virus preparations examined. As shown in figures III.2, III.3 and table III.1, the K_m values were comparable for either ApG primed or pGEM RNA primed transcription. The relative polymerase activities are depicted as incorporation of UMP or CMP (in moles) per mole of RNP. Because transcription was performed in a mixture of all 8 viral RNA segments, a mean vRNA length of 1700 nt (580 kD), packed with 85 NP monomers (4764 kD) and one polymerase complex (255 kD) was used as a basis to calculate specific polymerase activities. The concentration of purified RNPs was estimated from Coomassie-stained gels in comparison with BSA and optical density measurements at 260 nm. These values correlated reasonably well with 1 OD_{260} corresponding to 60 µg/ml RNP.

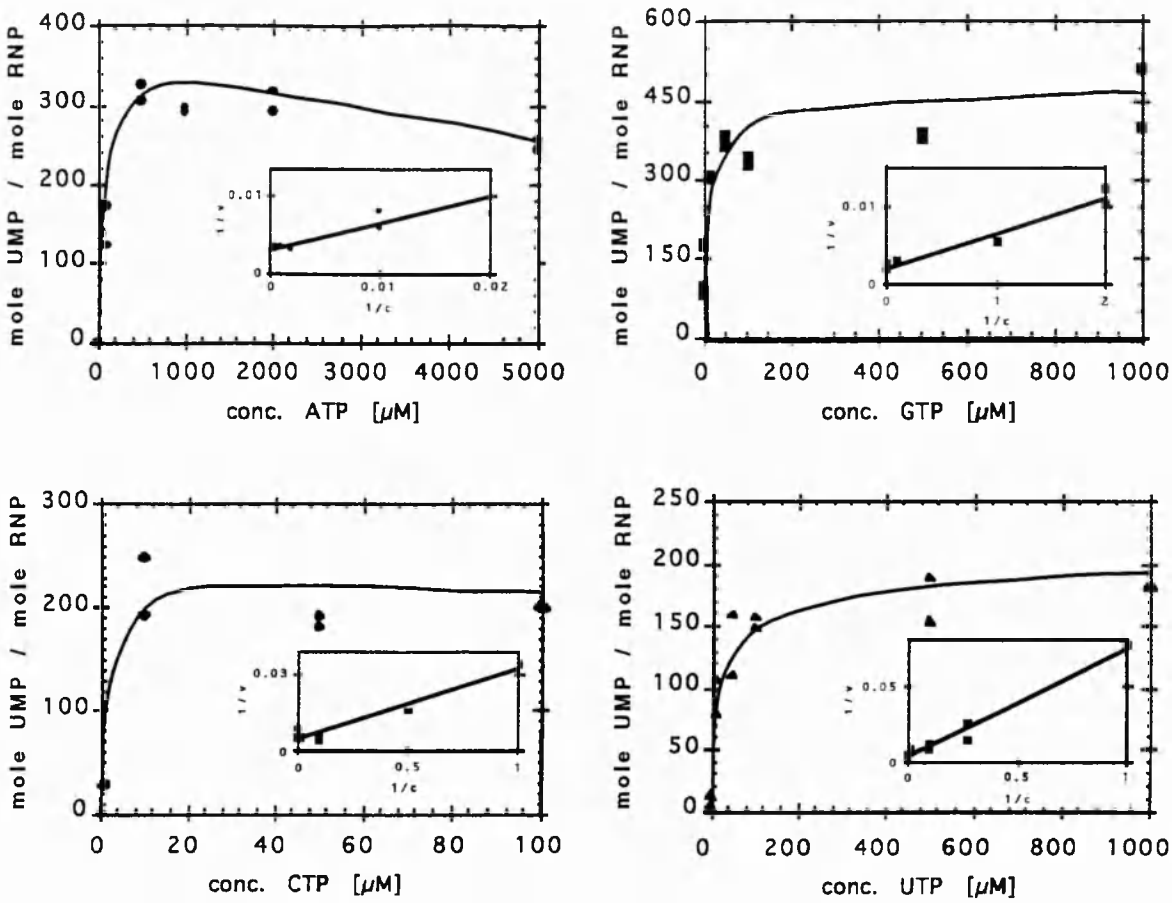


Figure III.2: RNP activity assay of ApG primed transcription. Standard transcription assays (see chapter IV) were performed with two unlabelled NTPs at a concentration of 500 μM , 400 μM ApG and 50 μM of one labelled NTP. The determination of [^{32}P]UMP or [^{32}P]CMP incorporation into trichloroacetic acid insoluble material was then performed in the presence of increasing concentrations of the fourth NTP as shown on the x-axis. Polymerase activity is depicted as relative incorporation of UMP or CMP (in moles) per mole of RNP in the assay. The insets show double reciprocal plots of the same experiment. The calculated K_m values are shown in table III.1.

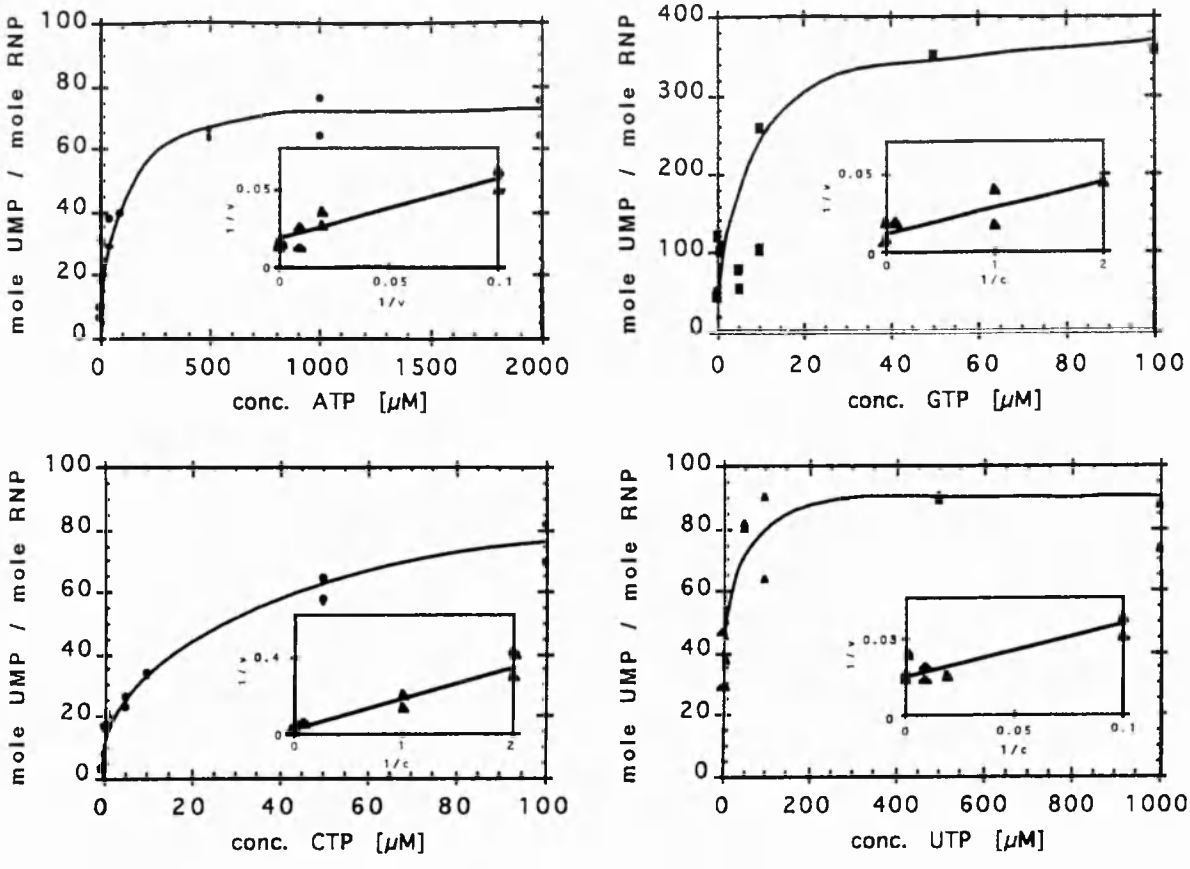


Figure III.3: RNP activity assay of capped pGEM primed transcription. Standard transcription assays (see chapter IV) were performed with two unlabelled NTPs at a concentration of 500 μM, 100 ng mRNA analog and 50 μM of one labelled NTP. The determination of [^{32}P]UMP or [^{32}P]CMP incorporation into trichloroacetic acid insoluble material was performed in the presence of increasing concentrations of the fourth NTP, as shown on the x-axis. Polymerase activity is depicted as relative incorporation of UMP (in moles) per mole of RNP in the assay. The insets show double reciprocal plots of the same experiment. The calculated K_m values are shown in table III.1.

Table III.1: The Km values in μM for the four substrates of the influenza virus RNPs as apparent from incorporation of radioactive nucleotides into acid insoluble material. The values for A/PR/8/34 were taken from figures III.2 and III.3, those for A/Vic/3/75 from (Stridh and Datema, 1984).

Primer	CTP	UTP	ATP	GTP
ApG				
K_m (A/PR/8/34)	6	10	120	2
K_m (A/Vic/3/75)*	22	14	153	14
pGEM RNA				
K_m (A/PR/8/34)	10	10	120	2
globin mRNA				
K_m (A/Vic/3/75)*	0.5	0.5	82	1.2

The K_m values for ATP were reproducibly found to be about 10-fold higher than those for the other nucleoside triphosphates. The same phenomenon has been observed with other influenza virus strains (table III.1), as well as with other RNA and DNA-dependent RNA polymerases (Krakow et al., 1976; Testa and Banerjee, 1979). This difference in K_m might simply reflect an adaptation of the viral polymerase to naturally occurring NTP concentrations in eucaryotic cells, where ATP levels are usually 7 to10-fold higher than other NTPs (Fairbanks et al., 1995; Stridh, 1983), but it is also possible that ATP plays an additional role other than that of a ribonucleoside triphosphate substrate used for chain elongation. ATP hydrolysis may be needed early in transcription as was observed with vesicular stomatitis virus (VSV) RNPs, another negative strand RNA virus (Perrault and McLear, 1984; Testa and Banerjee, 1979). VSV polymerase showed a high K_m for ATP during transcription initiation, but a significantly reduced K_m for ATP during elongation. One reason for this special requirement of ATP during transcription initiation of VSV polymerase may be the essential phosphorylation of the polymerase cofactor, the P protein, in this system (Chattopadhyay and Banerjee, 1987; Gao and Lenard, 1995). At the moment, it is not clear if there is a cofactor functionally equivalent to the VSV P protein in the influenza virus system.

T7 RNA polymerase, which is independent of ATP during transcription initiation, shows a K_m value for ATP similar to those for other NTPs (7-20 μM)(Chamberlin and Ryan, 1982; Osumi-Davis et al., 1992). Interestingly, it has been shown that T7 polymerase undergoes a conformational change early in transcription. The transcription initiation phase is characterized by a significantly higher K_m for GTP, the initiating nucleotide, and the production of abortive transcripts of 1-2 nucleotides length. Especially the production of short, abortive transcripts has been described for many RNA polymerases, that undergo conformational changes during RNA synthesis (Osumi-Davis et

al., 1992; Sousa et al., 1992). These examples show that RNA polymerases can go through a switch-like change in conformation and/or protein composition between transcription initiation and elongation phases.

Transcription by vesicular stomatitis virus RNPs has been analyzed in the presence of ATP analogs. When ATP was replaced by AMPPCP no activity was detectable in acid precipitation assays, although the analysis of reaction products revealed the formation of short, internally initiated, abortive transcripts (Perrault and McLear, 1984; Testa and Banerjee, 1979). Only when purified, ATP-preincubated RNPs were used, radioactive UMP incorporation into RNAs was observed, even when AMPPCP replaced ATP. Similar results have now been obtained with influenza virus RNPs. The ATP analogs AMPCPP and AMPPCP, which contain non-hydrolyzable α-β and β-γ bonds respectively, could not substitute for ATP in transcription assays (figure III.4A). AMPCPP serves as a control in this type of assay, because due to the blocked α-β bond it cannot be incorporated into any RNA transcripts. AMPPCP, on the other hand, possesses a hydrolyzable α-β bond and could in principle be used for incorporation during RNA synthesis, as has been observed with VSV polymerase during elongation. The fact that AMPPCP cannot substitute for ATP in influenza virus transcription may therefore indicate that, as in the VSV system, ATP β-γ bond hydrolysis is required for transcription initiation. However, it could also be that the influenza virus polymerase just does not have a high enough affinity for AMPPCP in order for it to support RNA synthesis. To get an idea for this affinity, in a second type of experiment elongation reactions were performed in the presence of 500 μM ATP, and the ability of ATP analogs to compete with ATP during transcription was measured. AMPCPP very strongly inhibited the influenza virus polymerase (figure III.4B) with half-maximal inhibition in the range of 100 μM AMPCPP. This suggests that AMPCPP can prevent ATP binding to the active site of the polymerase very efficiently and the polymerase appears to possess considerable affinity for the ATP analog. However, because AMPCPP has a blocked α-β bond, it cannot be incorporated into RNA and transcription is inhibited. There was much less transcription inhibition by AMPPCP and transcription could never be completely inhibited. Even the increase of ATP in the reaction led to a decrease in UMP incorporation during RNA synthesis (figure III.4B). The apparent differential inhibition of influenza virus specific transcription by the two ATP analogs and the fact that AMPPCP failed to completely inhibit influenza virus transcription, is consistent with incorporation of AMPPCP, but not AMPCPP, during elongation, although purely competitive inhibition by AMPPCP cannot be formally excluded from these results. The difference in inhibition between high concentrations of ATP and AMPPCP, as seen in figure III.4B, may be due to a slightly worse efficiency of AMPPCP α-β bond hydrolysis compared to ATP. Radioactively labelled ATP analogs might be used in the future to decisively answer the question of AMPPCP incorporation during elongation.

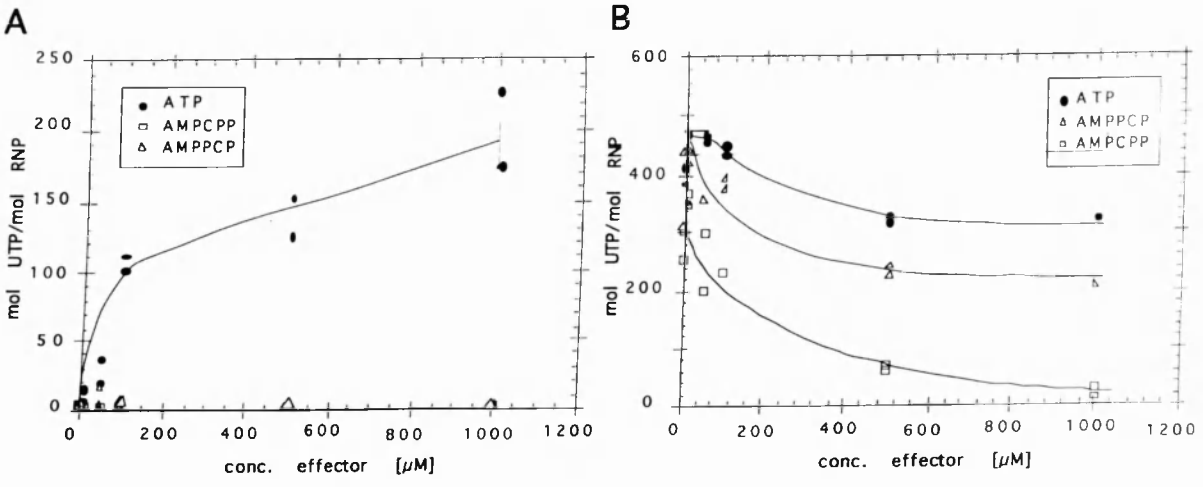


Figure III.4: ApG-primed *in vitro* transcription with influenza virus RNPs in the presence of increasing amounts of ATP or ATP analogs. (A) Reaction with 500 μM CTP and GTP, 50 μM labelled UTP (5 μCi) and up to 1 mM ATP or ATP analogs as indicated. (B), transcription reaction as in (A) including 500 μM ATP. Reactions were at 31°C for 1 h, followed by precipitation with 25% trichloroacetic acid.

Consistent with a role for ATP β-γ bond hydrolysis in influenza virus specific transcription, influenza virus RNPs were found to cosediment with an ATPase activity on glycerol gradients (figure III.5). ATPase activity was strongest with RNPs and RNP-ApG-C initiation complexes. When GTP and UTP were added to the reaction to promote the formation of elongation compexes, ADP production from ATP decreased. The same set of transcription reactions has been performed with saturating concentrations of unlabelled ATP, and the reaction products could be separated by reverse phase HPLC. ADP production was again evident and about 2-fold higher in initiation reactions compared to total transcription reactions including both initiation and elongation (figure III.6). The successful production of pre-initiated RNP-primer complexes will be required in order to determine, if ADP production occurs exclusively during transcription initiation.

Taken together, both the increased K_m for ATP in influenza virus transcription compared to the other NTPs, and the ATPase activity connected to influenza virus RNPs are consistent with a possible requirement of ATP β-γ bond hydrolysis early in transcription and support the idea of a conformational change between polymerase initiation and elongation complexes. In an attempt to identify the respective protein containing the ATPase activity in RNP preparations, we performed ATP-binding experiments with oxidized ATP as described in chapter II, II.2.1.3. However, the search for a specific ATP-binding protein in the RNP preparations has so far been unsuccessful, most likely due to methodological problems. Even PB1, which presumably contains the RNA polymerase domain and was meant to function as an internal ATP binding control protein, was not

specifically labelled by oxidized ATP during transcription elongation. In the future it might also be tried to separately purify initiation and elongation complexes and perform comparative studies of these two phases with regards to K_m values for NTPs. Such experiments will hopefully help to resolve the differences between the two hypothetical polymerase conformations.

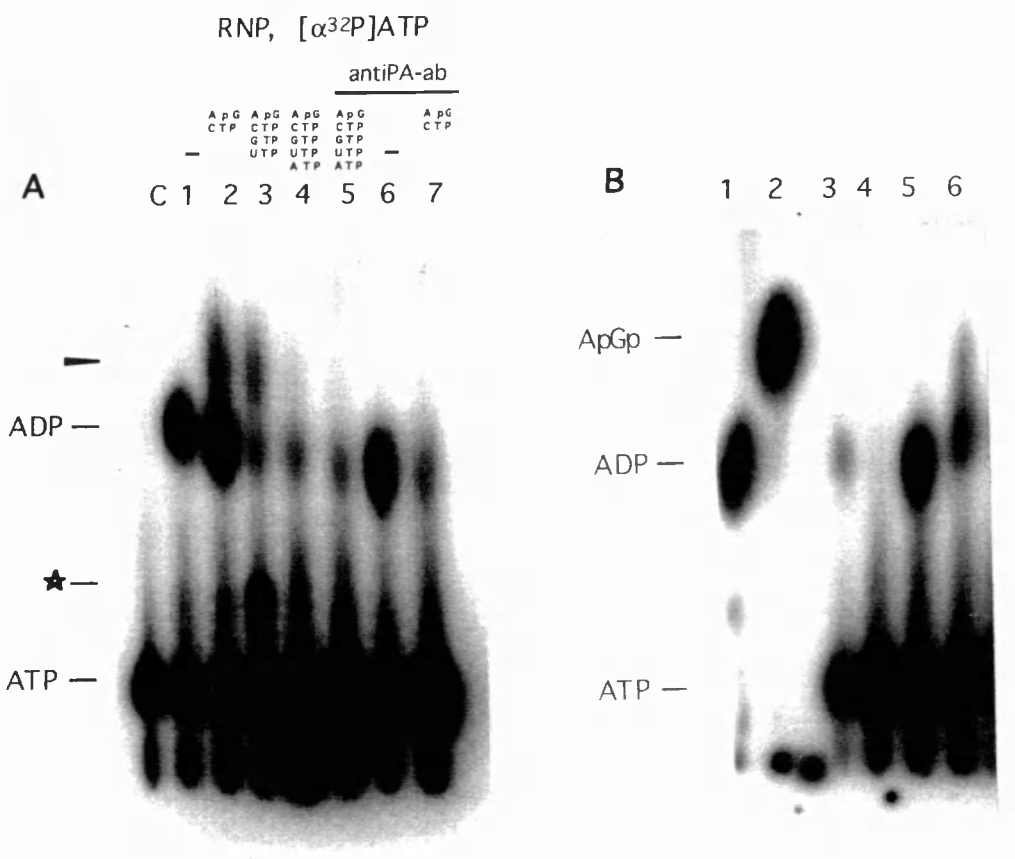


Figure III.5: ATPase activity of influenza virus RNPs. **(A)** 4 μg RNPs were incubated for 30 min at 31°C in transcription buffer with 5 μCi [α32P]ATP alone (lane **1** and **6**), plus additionally 400 μM ApG and 500 μM CTP (lane **2** and **7**), 400 μM ApG and 500 μM CTP, GTP, UTP (lane **3**), 400 μM ApG and 500 μM CTP, GTP, UTP, ATP (lane **4** and **5**). The RNPs were preincubated with 5 μl monoclonal anti-PA antibody (L35E12) in the reactions of lanes **5**, **6** and **7**. Lane "C" depicts [α32P]ATP incubated in transcription buffer. Reaction products were separated in 1 M LiCl, 0.5 M acetic acid on TLC plates. The positions of markers are indicated on the left. The arrowhead denotes a band, which is produced only in transcription initiation reactions. This band was not identified, but migrated similar to an ApGp marker. The star shows the position of a band, which is only produced in elongation reactions and may constitute a RNA product of limited size. **(B)** Labelled ADP marker was produced by incubation of 10 μCi [α32P]ATP with 0.2 μl T4-PNK and 100 ng DNA primer (lane **1**), labelled ApGp marker from incubation of 600 μM ApG with 0.5 μl T4 PNK with 10 μCi [γ32P]ATP (lane **2**). Lane **3** is a phosphatase reaction with limiting amounts of alkaline phosphatase, giving ATP and ADP marker. Lane **4**, incubation of [α32P]ATP in transcription buffer; lane **5**, incubation of [α32P]ATP with 4 μg RNPs in transcription buffer; Lane **6**, Transcription initiation reaction with RNPs, [α 32P]ATP, ApG and CTP.

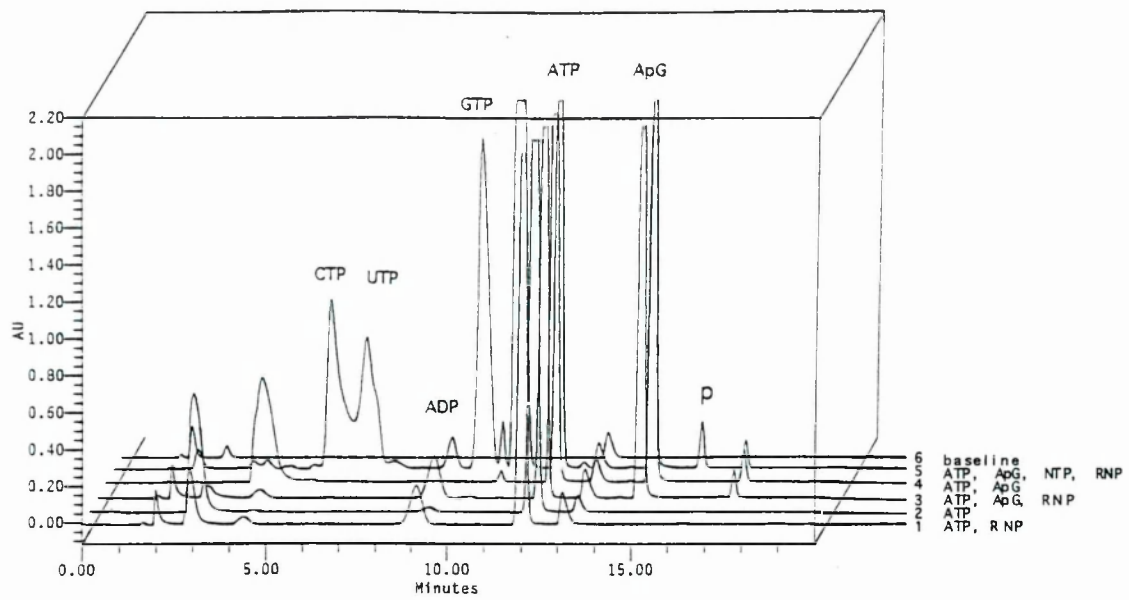


Figure III.6: HPLC separation of small nucleotides produced in RNP transcription reactions on a 2 ml Waters Symmetry C18 column. The x-axis denotes the retention time of nucleotides on the column as indicated, the y-axis relative absorption units measured at 260 nm. The separate experiments aligned along the z-axis show from front to back 1 h incubations at 31 °C in transcription buffer of (**1**) 2 µg RNPs with 1 mM ATP, (**2**) 1 mM ATP alone, (**3**) 2 µg RNPs with 1 mM ATP and 0.4 mM ApG primer, (**4**) 1 mM ATP and 0.4 mM ApG alone, (**5**) 2 µg RNPs with 1 mM ATP, 0.4 mM ApG and 0.5 mM each CTP, UTP, GTP. The experiment shows ADP production only in the presence of RNPs and a 50% reduction in ADP production in a complete transcription reaction (5), compared to free RNPs (1) and RNP-primer complexes (3). "p" denotes an additional peak, which only appears after transcription initiation and presumably constitutes a small RNA product. The chromatography was performed in 0.1 M potassium phosphate, pH 6.5, 5 mM tetrabutylammoniumhydroxide (TBA). Bound nucleotides were eluted with a 0-50% methanol gradient in the same buffer (20 min at 1 ml/min).

III.4.2 Transcription reactions primed with capped RNA molecules

III.4.2.1 Characterization of transcription initiation and elongation reactions

As shown in Figure III.7A, a reaction including viral RNPs, pGEM RNA and labelled CTP leads to the production of a single 12 nt RNA product of transcription initiation (lane 12, "initiation reaction"), as also shown schematically in figure III.1, page 84. In the additional presence of unlabelled ATP, GTP and UTP, elongation to full-length RNAs is observed, which stay near the top of 20% urea-PAGE gels (figure III.7A, lanes 13-16, "elongation reactions"). In this, as well as in the following experiments, there was always some unspecific residual radioactivity, that stayed in the wells of the 20% acrylamide gels. The origin of this radioactivity was not explicitely determined, but most likely reflects insoluble precipitates formed in the transcription reactions after addition of the

formarnide loading buffer and heating to 95°C before loading. Therefore, only bands, that had at least slightly entered the gel were considered as specific and included in quantitative measurements. Lanes 1-8 in figure II.7A show a titration of GTP into a transcription initiation reaction with pGEM RNA and labelled CTP. At low GTP concentrations, no GTP, but only the labelled CTP is exclusively added to the primer RNA and a band corresponding to G11+1nt is visible (lanes 1 and 2). A band of G11+2nt is first observed at considerable excess of GTP (40-fold) in the reaction (lane 3). As shown on figure III.1, page 84, this second band can either be formed by transcription initiation with CTP and incorporation of GTP as the second nucleotide, or by initiation with GTP and incorporation of CTP as the second nucleotide. At very high GTP concentrations, G11+3nt products are observed (lanes 7 and 8), that presumably stem from initiation with GTP and subsequent addition of CTP and GTP (see also figure III.1). According to this experiment the first nucleotide, CTP, seems to be more efficiently added to the primer than the second one, GTP. The band intensities decrease with increasing GTP in the titration, because of competition between GTP and labelled CTP for transcription initiation.

Even in the presence of all four nucleotide triphosphates, significant amounts of doublet bands of 12 and 13 nucleotide lengths were still produced together with long transcripts (figure III.7A, lanes 14, 15). These bands are most likely products of abortive transcription initiation, and they are especially strong at lower nucleotide concentrations. The production of either very short, distinct, abortive initation transcripts or large transcripts in the absence of intermediate size RNAs can be considered as an indication for a limiting step early in transcription, e.g. a conformational change between transcription initiation and elongation complexes, as has been discussed before for other RNA polymerases (Jacob et al., 1994; Kubori and Shimamoto, 1996; Sousa et al., 1992). With influenza virus RNPs, the incorporation of the first ATP, which has to be used as the second or third nucleotide, appears to be a rate limiting step during early transcription. Consistent with this notion, we never succeeded in efficiently producing G11+2nt RNAs in the presence of CTP and ATP or G11+3 RNAs in the presence of CTP, GTP and ATP. Likewise we never captured G11+3, +4 and +5 RNAs. Instead, with CTP, GTP and ATP present, the RNPs either produced abortive transcripts lacking adenosine (A) residues or RNAs of 20-30 nucleotides length, which are complete transcripts halted at the position, where the first U residue were to be incorporated into mRNA (see figure III.9, page 97 and scheme in III.1, page 84). All these observations correlate well with the previously discussed results of transcription reactions in the presence of ATP analogs and the ATPase activity of RNPs and support a model of the influenza virus polymerase complex going through a measurable change between transcription initiation and elongation phases.

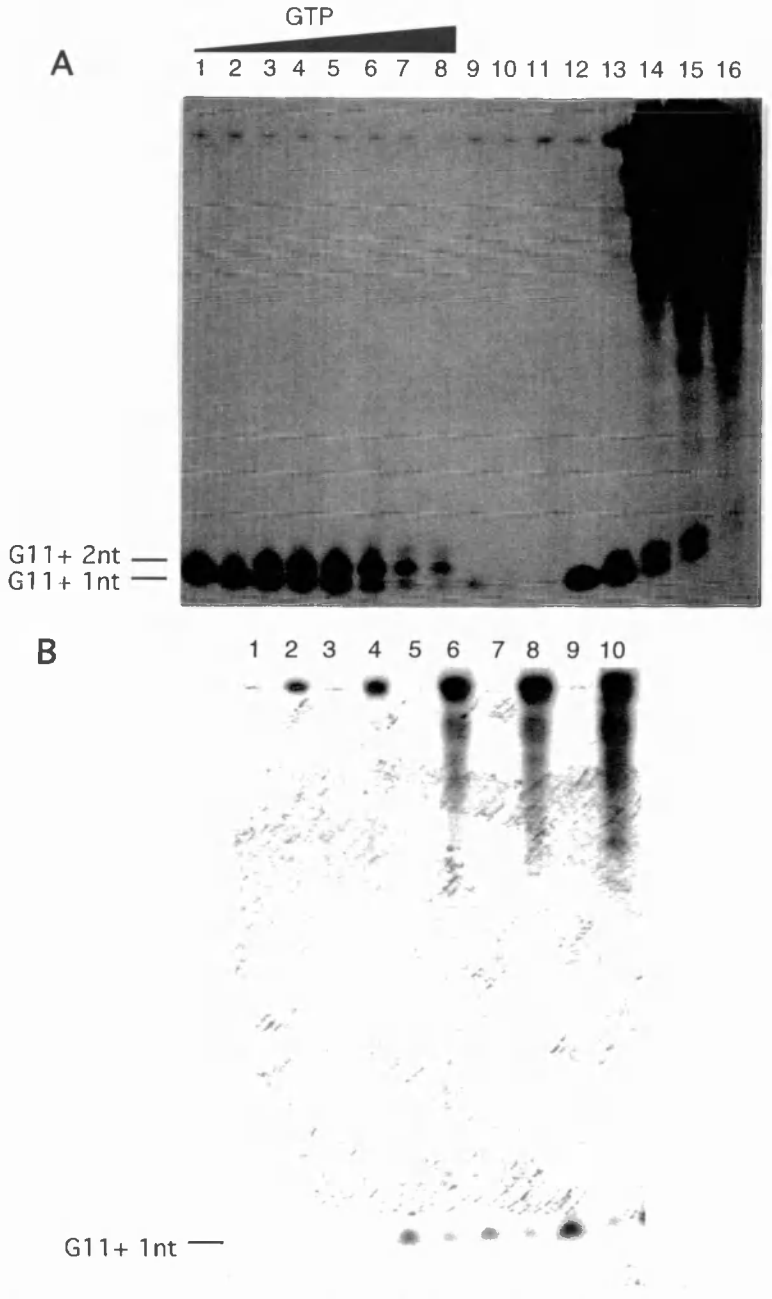


Figure III.7: *In vitro* transcription activity of influenza virus RNP. 3 μg purified RNPs were incubated with 0.25 μM [α^{32}P]-CTP (5 μCi) and 0.1 μg capped pGEM RNA in transcription buffer to which other nucleotides were added at variable concentrations. The reaction products were analyzed on 20% urea-PAGE gels. **(A)**, effect of GTP on transcription initiation. Both GTP and CTP can be incorporated during transcription initiation. Lanes, **1-8**, transcription reaction with increasing concentrations of GTP (0.2, 1, 5, 10, 16, 50, 100, 200 μM respectively); lane **9** and **10**, addition of 8 μM 2'-FdCTP without (**9**) or with (**10**) 1 μM ATP, GTP, UTP to the standard initiation reaction; lane **11**, addition of 8 μM CTP to the initiation reaction; lane **12**, basic transcription initiation reaction with only [α^{32}P]-CTP, lanes **13-16**, addition of ATP-GTP-UTP-mix at 1, 5, 10 and 100 μM. **(B)** Transcription inhibition by cap analogs. Reactions have been done and loaded in pairs. Uneven numbers are initiation reactions with 0.25 μM [α^{32}P]-CTP (5 μCi) as only NTP. Even numbers are elongation reactions with 100 μM ATP, GTP, UTP in addition to the label. These reactions were performed in the presence of 0.2 mM cap I (m7GpppGm) (lanes **1**, **2**), 0.2 mM cap 0 (m7GpppG) (lanes **3**, **4**), 0.48 mM m7GTP (lane **5**, **6**), 0.87 mM m7GDP (lanes **7**, **8**). Lanes **9**, **10** are control reactions in the absence of inhibition.

No transcription initiation was observed with uncapped pGEM RNA, confirming the results of Chung et al. (1994) and illustrating the specificity of the transcription reaction as well as the need for a cap structure. Transcription initiation was specifically inhibited by cap analogs (figure III.7B). Consistent with the results of Blaas et al. (1982), inhibiting activity of cap anlogs in influenza virus specific transcription decreased in the series cap I (m7GpppGm) > cap 0 (m7GpppG) >> m7GTP > m7GDP. These results demonstrate, that the present transcription assay really measures cap-dependent transcription initiation.

III.4.2.2 Transcription inhibition by 2'-deoxy-2'-fluoroguanosine triphosphate

The above described transcription assay was used to study the mechanism of influenza virus inhibition by 2'-deoxy-2'-fluororibonucleotides. The addition of 40 μM 2'-deoxy-2'-fluoroguanosine-triphosphate (2'-FdGTP) to the transcription initiation mixture resulted in the appearance of two products of 12 and 13 nucleotides respectively (figure III.8, lane 5). The same doublet of bands was produced, when 40 μM GTP replaced 2'-FdGTP in the reaction (lane 6). This result strongly suggests that the nucleotide analog 2'-FdGTP is recognized by the influenza virus polymerase and can be incorporated into specific, capped mRNAs during the initiation of transcription, with an efficiency similar to that of GTP. The doublet of bands showed 40-50% reduced incorporation of total radioactivity compared to a standard initiation reaction in the absence of effector (lane 1 or 3), according to band density scanning of autoradiographs. This reduction in band intensity is most likely due to competition between unlabelled GTP and labelled CTP for being used during transcription initiation (see above in III.4.1.1). From the comparable reduction of total radioactivity it can be concluded that both 2'-FdGTP and GTP compete with comparable efficiency with labelled CTP at transcription initiation. The similarity of the doublet bands produced in the presence of GTP or 2'-FdGTP implies that no significant inhibition of the influenza virus polymerase complex occurs prior to the incorporation of the second nucleotide during mRNA synthesis. Specific inhibition of transcription became apparent only in elongation reactions with all four NTPs and including 2'-FdGTP. The production of long transcripts was significantly reduced in the presence of 2'-FdGTP compared to control reactions with GTP (figure III.8, lanes 7 and 8). The lower intensities of the transcription elongation products in the presence of 2'-FdGTP, together with the fact that 2'-FdGTP can be incorporated by the influenza virus polymerase into nascent mRNAs, suggested that the nucleotide analog may function as a chain terminator after incorporation.

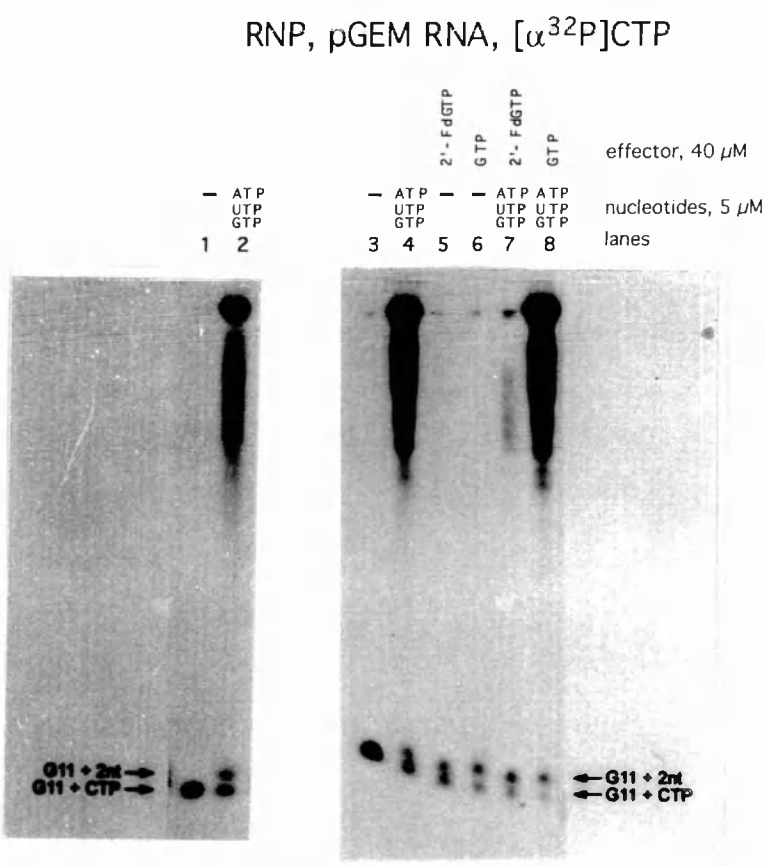


Figure III.8: *In vitro* transcription in the presence of the nucleotide analog 2'-deoxy-2'fluoro-guanosine-triphosphate (2'-FdGTP). The figure shows two separate experiments for the standard initiation (lanes 1 and 3) and elongation (lanes 2 and 4) reactions. All reactions were performed with 3 μg RNPs, 0.1 μg pGEM RNA and 0.25 μM [α^{32}P]-CTP (5 μCi). The reactions loaded in lanes 2, 4, 7 and 8 additionally contained 5 μM each ATP, UTP and GTP. Either 40 μM 2'-FdGTP or 40 μM GTP were added to the reactions as indicated above the lanes (lanes 5-8). The reaction products were analyzed by autoradiography of 20% urea-PAGE gels.

Limited elongation reactions in the absence of UTP also support the idea of a chain termination mechanism of transcription inhibition by the fluorinated nucleotide analogs. In the absence of UTP, the polymerase protein can only elongate the nascent mRNAs until the first adenine residue on the vRNA template. According to the sequence variation in the different RNA segments, mRNA molecules of 21-31 nucleotides length can be produced in transcription reactions with pGEM RNA primer, viral RNPs, CTP, ATP and GTP (figure III.9, lane 2; figure III.9 shows two different exposure times of the same gel in order to resolve initiation bands as well as limited elongation bands). A parallel set of limited elongation reactions like the ones in figure III.9, lanes 1 and 2 were preincubated for the same time as in lanes 1 and 2 and then provided with additional NTPs (ATP, UTP, GTP) to allow complete elongation. Lane 4 shows that in this way large amounts of long transcripts can be produced from such preincubated limited elongation reactions. In the presence of 2'-FdGTP, replacing GTP during limited elongation, virtually no RNAs between 21-31 nucleotides are synthesized (lane 1). Apart from the 12 nt and 13 nt transcripts only a very low intensity band of about 19 nt length is synthesized in the presence of 2'-FdGTP, labelled CTP and ATP (lane 1). Transcription in lane 1 compared to lane 2 appears to be blocked at G11+2nt, i.e. at the position, where the first G residue is incorporated into the nascent RNA (see sequence scheme in figure III.1, page 4). This result supports the concept that incorporation of 2'-FdGTP early in transcription blocks the elongation of mRNAs by the influenza virus polymerase.

However, both prematurely terminated bands of 13 nt and 19 nt diminish after the addition of further nucleotides, during elongation (lane 3), suggesting that the apparent block of transcription by 2'-FdGTP incorporation is reversible. Consistent with the reversibility of inhibition by 2'-FdGTP was the observation that RNA synthesis was strongly inhibited in reactions with viral RNPs, ApG primer, ATP, CTP labelled UTP and 2'-FdGTP. But even after extensive preincubation of this reaction in order to produce maximal amounts of blocked transcripts, transcription efficiently resumed, when large amounts of GTP were added to the preincubated reactions (Tisdale et al., 1995). These results suggest that the influenza virus RNPs are capable of regenerating blocked transcripts. Before going into further discussion of this unusual activity for a RNA polymerase, the following paragraph will show examples of transcription reactions in the presence of another fluorinated nucleotide analog, 2'-FdCTP. The observations with the cytidine analog support the model of transcription inhibition caused by compound incorporation and elongation block. Finally, it will be shown that the above results were also confirmed by performing the transcription reactions with globin mRNA as a primer of influenza virus transcription instead of the pGEM RNA.

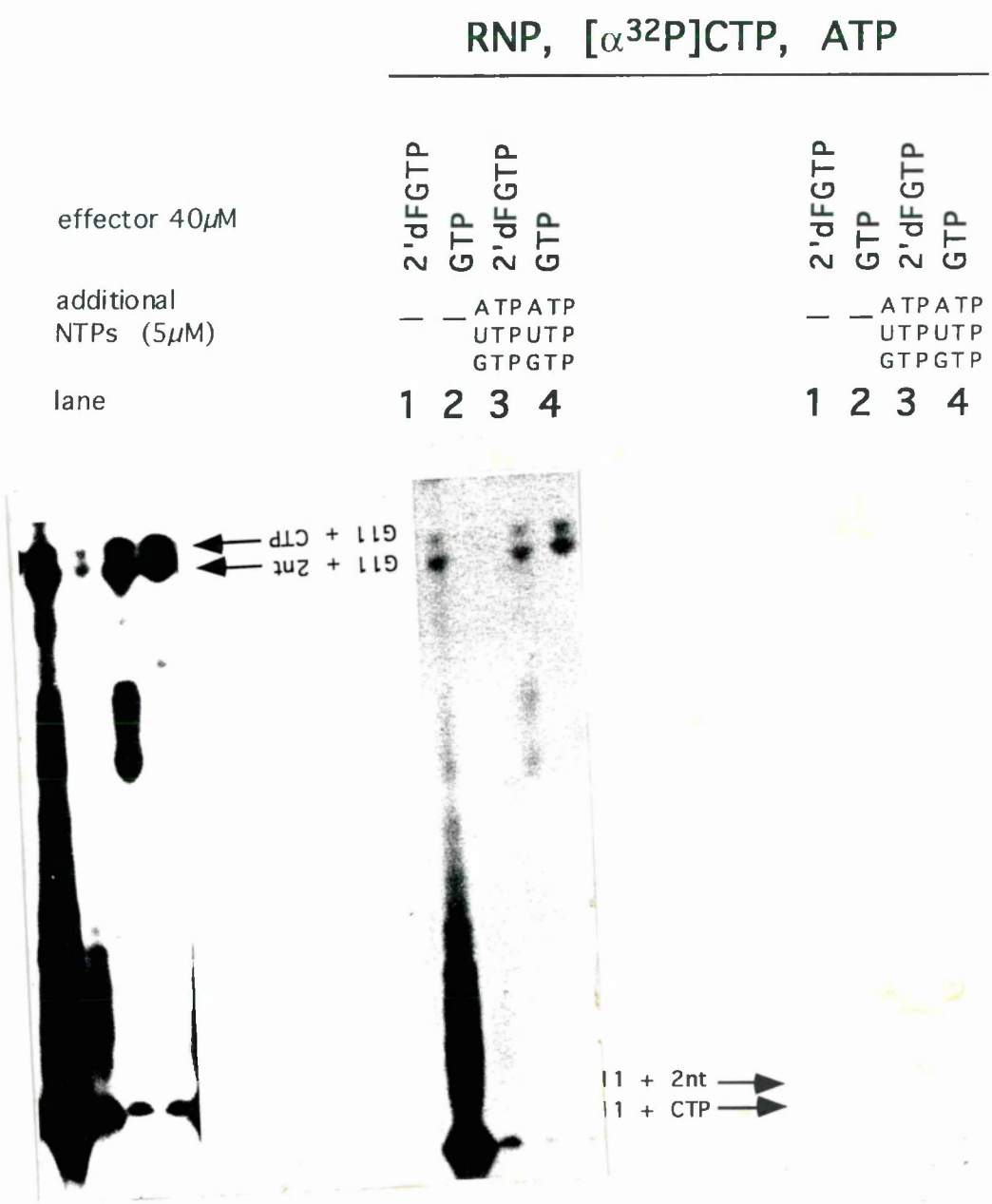
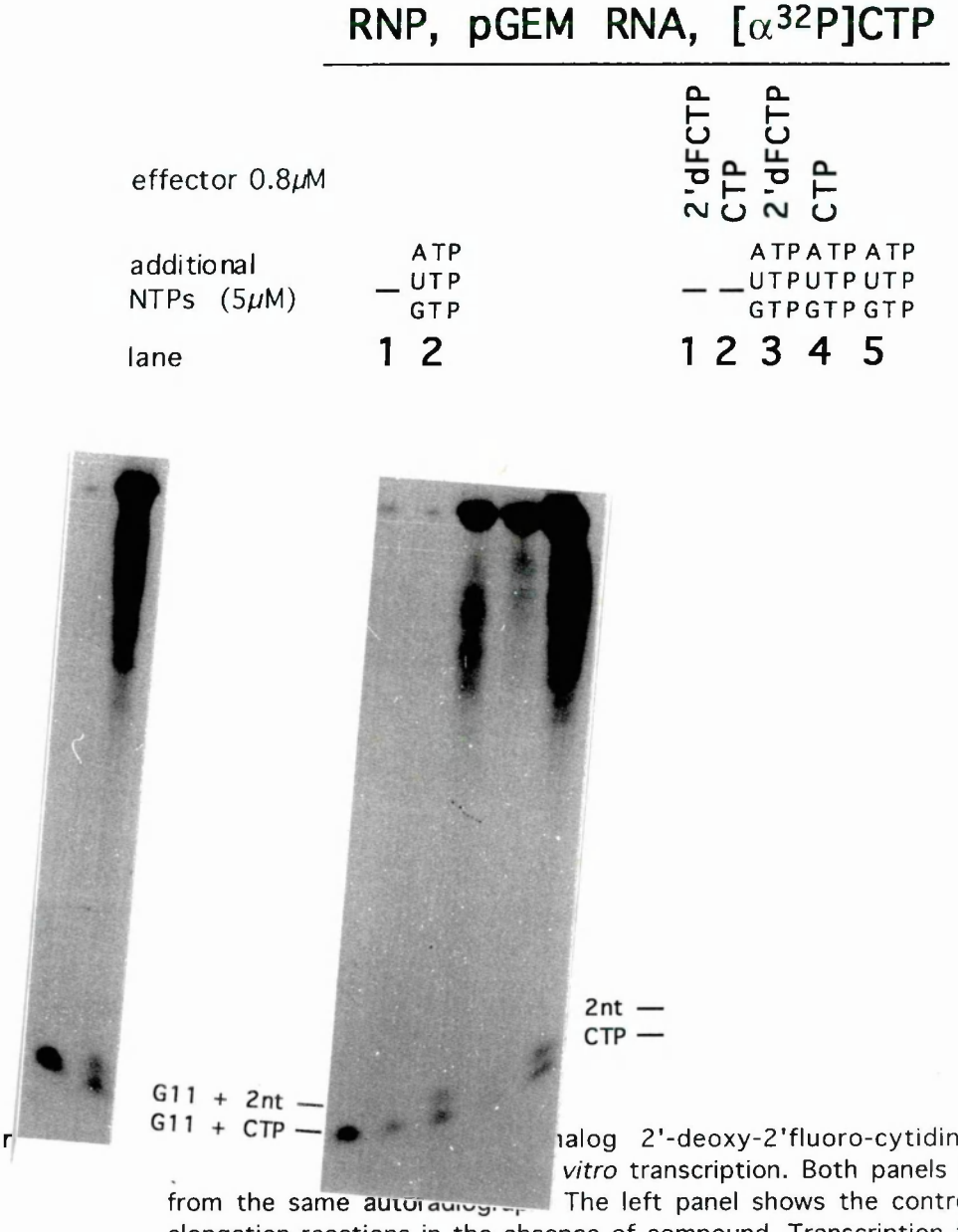


Figure III.9: *In vitro* transcription in the presence of the nucleotide analog 2'-deoxy-2'fluoro-guanosine-triphosphate (2'-FdGTP). The figure shows two different exposure times of the same experiment. 2'-FdGTP prevents limited elongation and blocks transcripts at the position G11+2nt, where it is presumably incorporated (see scheme in figure III.1, page 84). Transcription from 3 μg RNPs was primed with 0.1 μg pGEM RNA and 0.25 μM [α^{32}P]-CTP (5 μCi) and 5 μM ATP in the additional presence of either 40 μM 2' F-dGTP (lane 1) or 40 μM GTP (lane 2). To switch from limited to complete elongation, 5 μM ATP, UTP and GTP were added to **preincubated** limited elongation reactions. Lanes 3 and 4 show the products of complete elongation when the samples of lanes 1 and 2 were supplemented with 5 μM ATP, UTP and GTP. The reaction products were analyzed by autoradiography of 20% urea-PAGE gels.

III.4.2.3 Transcription inhibition by 2'-deoxy-2'-fluorocytidine triphosphate

Similar experiments as described above for 2'-FdGTP were performed with the nucleotide analog 2'-deoxy-2'-fluorocytidine-triphosphate (2'dFCTP). Figure III.10 shows on the left panel one more standard transcription initiation and one elongation reaction. The right panel of figure III.10 shows that prematurely terminated transcripts accumulate during elongation in the presence of 2'-FdCTP (right panel, lane 3), but not when 2'-FdCTP was replaced by CTP (lane 4; both panels of figure III.10 were cut out from the same autoradiograph). These reactions were done with very low amounts of inhibitor (0.8 μM) in order to avoid too strong a competition of the unlabelled CTP analogs with the labelled CTP in these transcription reactions. This is the reason that, apart from the apparent inhibitory effect on elongation, there were also reasonable amounts of long transcripts produced in lane 3. In the experiments with 2'-FdCTP the competition with the labelled CTP reduced the overall signal obtained in the reactions in the presence of the effectors. CTP appeared to be a 3-fold better competitor than 2'-FdCTP for the labelled CTP, as determined by density scanning of the corresponding bands on the autoradiograms. In figure III.10, the formation of the labelled G11+1nt transcription initiation band was reduced to 20% of the control in the presence of cold CTP as effector control (compare right panel, lane 2 with left panel, lane 1), but the G11+1nt initiation band was only reduced to 60% of the control with 2'-FdCTP (compare right panel, lane 1 with left panel, lane 1). This result strongly suggests that, as before with 2'-FdGTP, transcription initiation and earlier steps (cap-binding, endonuclease activity) are not inhibited by the fluorinated nucleotide analogs. Both guanosine and cytidine analogs appear to mainly inhibit transcription elongation.

In the following, preincubation of RNPs with 2'-FdCTP was used to address again the question of reversibility of transcription inhibition, that had before become apparent with the guanosine analog 2'-FdGTP. The experiments shown above in paragraph III.4.2.2 established that 2'-FdGTP could be incorporated into transcripts early in transcription, and blocked transcripts corresponding to G11+2nt accumulated in the reactions (figure III.9). We also obtained evidence that 2'-FdCTP could be incorporated into transcripts in a similar manner, an exemplary gel of which will be shown and discussed further below (figure III.13, lanes 17-20). From a time course reaction it appeared that a 30 min incubation was sufficient for maximal CMP incorporation during transcription initiation to form maximal amounts of G11+1nt initiation products. We therefore expected to produce significant amounts of blocked G11+1nt products when incubating RNP-primer complexes with only 2'-FdCTP. Interestingly, the preincubation for 30 min of RNP-primer complexes with only 2'-FdCTP before adding labelled CTP for another 30 min had only a minor effect on the production of the G11+1nt RNA product compared to reactions, where both compound and label were added simultaneously (figure III.11, compare lanes 1 and 5).



Figur alog 2'-deoxy-2'fluoro-cytidine-triphosphate *vitro* transcription. Both panels represent lanes from the same autoradiograph. The left panel shows the control initiation and elongation reactions in the absence of compound. Transcription from 3 μg RNPs was primed with 0.1 μg pGEM RNA and 0.25 μM [α^{32}P]-CTP (5 μCi) in the additional presence of either 5 μM ATP, GTP and UTP, 0.8 μM 2'-FdCTP or 0.8 μM CTP, as indicated above each lane. In the presence of the nucleotide analog, the transcript lengths are reduced during elongation, consistent with the idea of a transcription block or low efficiency elongation after compound incorporation. There are still long transcripts produced in the presence of compound, because the reaction was done at the lower detection level of the inhibitory effect. The reaction products were analyzed by autoradiography of a 20% urea-PAGE gel.

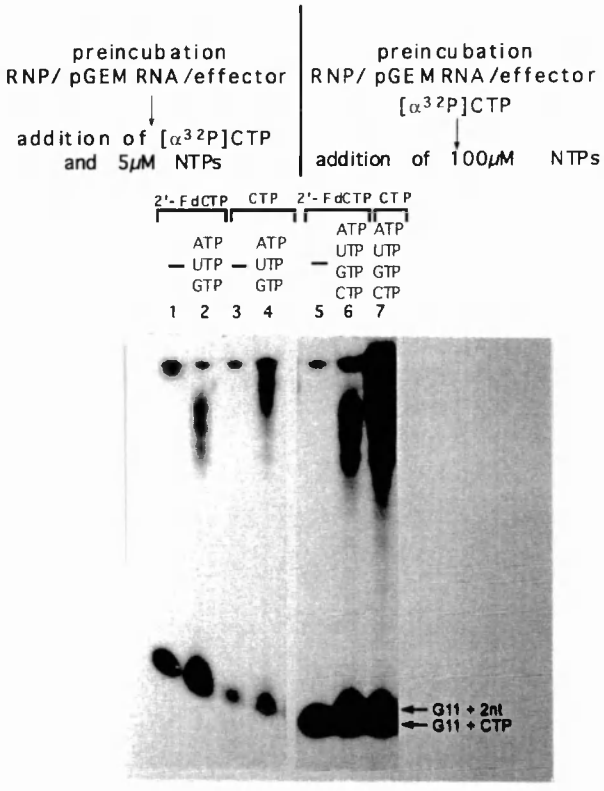


Figure III.11: *In vitro* transcription in the presence of the nucleotide analog 2'-deoxy-2'fluoro-cytidine-triphosphate (2'-FdCTP). The reaction conditions were as described in figure III.10. Lanes 1-4, preincubation of RNP, pGEM RNA and 0.8 μM 2'-FdCTP (1, 2) or CTP (3, 4) for 30 min, then addition of 0.25 μM [α^{32}P]-CTP (5 μCi) (lanes 1, 3) or 0.25 μM [α^{32}P]CTP plus 5 μM ATP, UTP, GTP (lanes 2, 4) for another 30 min. Lane 5, incubation of RNP, pGEM RNA, 0.8 μM 2'-FdCTP and 0.25 μM [α^{32}P]CTP for 30 min; lane 6, preincubation as in lane 5, then addition of 100 μM NTPs and incubation for further 30 min; lane 7 reaction as described for lane 6 with 0.8 μM CTP replacing 2'-FdCTP.

To account for this observation, either there was no incorporation and block of transcription at initiation (G11+1nt production), or the RNPs regenerated G11 primers for de novo transcription from blocked G11+1nt transcripts, as seen before after 2'-FdGMP incorporation. An inhibitory effect of the compound still became clearly apparent during elongation, when the length of the transcripts was again reduced in the presence of 2'-FdCTP compared to the control reaction with CTP replacing the compound (figure III.11, compare lanes 2 and 4). In both CTP control reactions the G11+1nt bands were of lower intensity than in the corresponding compound reactions due to the previously mentioned threefold higher competition of CTP compared to 2'-FdCTP for labelled CTP. At this stage, the experiments with the cytidine analog could not yet exclude a purely competitive inhibition mechanism by 2'-FdCTP, although this would mean that the cytidine compound behaved differently than the guanosine analog. Pulse-chase experiments were used to further study the mechanism of transcription inhibition by 2'-FdCTP in this system. The RNP-primer complexes were first preincubated with labelled

CTP and 2'-FdCTP (figure III.11, lane 5), before being chased with 100 μM ATP, GTP, CTP and UTP, i.e. a 1000-fold excess of the nucleotides over the label CTP and 125-fold excess over 2'-FdCTP (figure III.11, lane 6). The pattern of transcription inhibition was similar to the previously shown reactions. Even in the presence of excess nucleotides in the reaction the amount and the length of the transcripts were reduced with 2'-FdCTP compared to control reactions with CTP replacing 2'-FdCTP (compare lanes 6 and 7). If inhibition were exclusively at transcription initiation, then in pulse-chase experiments the fraction of non-blocked transcripts should be completely elongated. The preferential formation of shortened transcripts in the chasing period implies an influence of the inhibitor during elongation even at very low concentrations relative to the homologous NTP. The fact that an inhibitory effect, i.e. the reduction of transcript length, is still measurable under those conditions supports the chain blocking mechanism of transcription inhibition rather than a purely competitive mechanism. The formation of a labelled, blocked G11+1nt product could not be explicitely shown without having labelled 2'-FdCTP available. However, the observation that the fluorinated compound could be incorporated during early transcription with another primer RNA (discussed below, figure III.13, lanes 17-20), together with the absence of significant inhibition of G11+1nt RNA production after preincubation of RNPs with 2'-F-dCTP can be considered as a further evidence for a regeneration activity, i.e. the ability of the RNP to take presumably blocked G11+1nt products and use them to regenerate G11 primers.

The results so far suggested that inhibition of influenza virus specific transcription by the nucleotide analogs 2'-FdGTP and 2'-FdCTP followed a similar mechanism and occurred mainly, if not exclusively, during elongation, most likely due to chain termination after incorporation into nascent RNAs. This mechanism is also the most plausible to account for the remaining inhibitory effect even during chasing with large excess homologous nucleotides.

III.4.3 Influenza virus transcription reactions primed with globin mRNA

The following experiments, using globin mRNA as primers in combination with labelled GTP instead of labelled CTP, were used to study the impact of the nucleotide analogs on influenza virus transcription in yet another system. One interest was to use a different, labelled nucleotide in order to exclude that the previously apparent mechanism of transcription inhibition was influenced by the competition of the compounds with the label. Also it was possible to directly visualize the incorporation of unlabelled 2'-FdCTP during early transcription into primers initiated with GTP, as seen before with 2'-FdGTP in the pGEM system initiated with CTP. Rabbit globin mRNA is a mixture of α- and β-globin mRNAs. Plotch et al. (1981) have shown that influenza virus endonuclease preferentially cleaves β-globin mRNA in a mixture of globin mRNAs and produces a capped primer RNA of 13 nucleotides, called G13 (figure III.12).

Figure III.12: Schematic representation of the presumed primer-vRNA interactions and the nucleotides incorporated by influenza virus RNPs during transcription initiation with globin mRNA according to Plotch et al. (1981). β-globin is the mRNA, that is preferentially used for transcription initiation in the globin mRNA mixture. The influenza virus endonuclease cleaves mainly 3' of G13 (A), and the first nucleotide incorporated in vitro is a GTP (B). In the presence of elevated concentrations of GTP as only NTP, several G residues can be incorporated into the nascent RNA (B). In the presence of both GTP and CTP, only one GTP and one CTP are incorporated (C). When α-globin is used as endonuclease substrate, it is cleaved 3' of G10 (D) and initiated exclusively with CTP (E).

The influenza virus polymerase initiates transcription from β-globin mRNA derived primers preferentially with GTP. In the presence of GTP as the only nucleotide, several G residues can be added to the primer sequence in a stuttering mechanism (Plotch et al., 1981). A similar phenomenon was observed in the present study: when globin mRNA was incubated with viral RNPs and labelled GTP, two bands migrating at 14 nt and 16 nt were generated, presumably corresponding to G13+1nt and G13+3nt (figure III.13, lane 3). Both bands disappeared when CTP was available and were replaced by a band of intermediate mobility (15 nt, i.e. G13+2nt; figure III.13, lane 2). This suggested that in the presence of only GTP one or three G residues were added to the primer, whereas with both GTP and CTP present, only one of each nucleotide was incorporated (see scheme in figure III.12). Long transcripts were efficiently produced in the presence of all four nucleotides (lane 4).

Similar to CTP (lane 2), 2'-FdCTP induced the formation of a single 15 nt band in the presence of viral RNPs and labelled GTP (lane 17). This clearly demonstrated that 2'-FdCTP, as described above for 2'-FdGTP, could be incorporated into nascent mRNAs by the influenza virus polymerase with an efficiency similar to CTP. The band intensity was again reduced compared to the standard reaction (compare lanes 2 and 17; 2 and 19), presumably due to a competing side reaction, similar to the one described above. It is known, that α-globin mRNA can serve as an endonuclease substrate, when only CTP is present. In this case, α-globin is cleaved 10 nt downstream of the cap structure and transcription initiates by incorporation of a C residue (Plotch et al., 1981)(see scheme, figure III.12). This alternative initiation from α-globin-derived primer with CMP incorporation might compete with transcription initiation from β-globin using labelled GTP. This competition appears to be strong enough that in the presence of elevated CTP levels (100 μM), relative to labelled GTP (0.25 μM), transcription initiation from β-globin is strongly reduced (lane 19). Lanes 18 and 20 illustrate the inhibition of elongation by 2'-FdCTP compared to CTP. As before, both the amount and the lengths of the transcripts are reduced.

In the globin mRNA system, the nucleotide analogs had again no detectable effect on transcription initiation, but significantly reduced the amount and lengths of RNAs synthesized during elongation relative to control reactions with CTP replacing 2'-FdCTP and GTP replacing 2'-FdGTP respectively. In figure III.13 lanes 5-10 are transcription reactions with 0.8 μM 2'-FdGTP or GTP, lanes 11-16 with 4 μM 2'-FdGTP or GTP respectively. Lanes 5, 7, 11 and 13 show no significant difference during transcription initiation in the presence of either nucleotide analog or GTP. However, incorporation of up to 4 G residues during transcription initiation is only seen with 4 μM GTP alone. (lane 13). The presence of 2'-FdGTP does not inhibit the formation of the G13+3nt initiation product, that is formed in the presence of only GTP. (compare lanes 5 and 11 with lane 3). At higher concentrations of 2'-FdGTP or GTP (lanes 11-16), competition of GTP for the labelled GTP (0.25 μM) became apparent, as seen before with CTP in the pGEM-RNA system with labelled CTP. The competition between GTP or 2'-FdGTP for labelled GTP led to an overall decrease in the intensity of transcription products. As observed before, the unlabelled nucleotide was a stronger competitor for the incorporation of homologous label than the nucleotide analog. Therefore, the intensity of the initiation bands was lower in the GTP reaction (lane 13) than in the 2'-FdGTP reaction (lane 11). Nevertheless, only 2'-FdGTP prevented the polymerase from efficiently synthesizing long RNAs (lanes 12 and 14). During elongation, both the amount and the lengths of the synthesized RNAs were reduced. The 30 min preincubation of RNP-primer complexes with 2'-FdGTP, in order to incease the number of RNAs initiated with the nucleotide analog, did not change the pattern of RNA synthesis during elongation and did especially not further inhibit transcription (compare for example lanes 15 and 16 after preincubation with lanes 12 and 14 without preincubation).

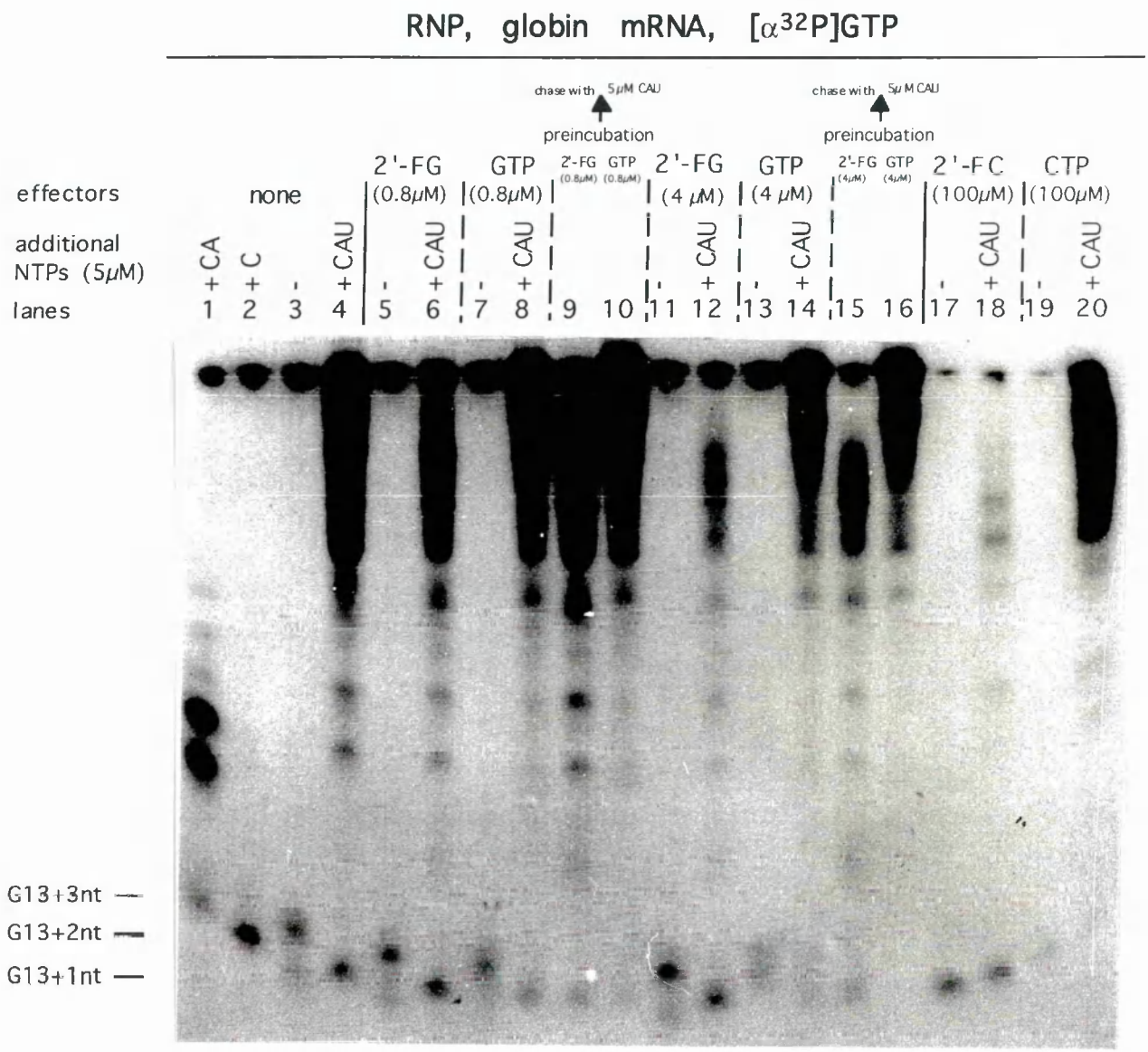


Figure III.13: *In vitro* transcription in the presence of nucleotide analogs and with 0,1 μg globin mRNA as primer. For standard reactions 0.25 μM [α^{32}P]-GTP (5 μCi) was used alone (lane **3**, initiation), or in combination with 5 μM CTP and ATP (lane **1**, limited elongation), 5 μM CTP (lane **2**, incorporation of 2nd nucleotide), 5 μM CTP, ATP and UTP (lane **4**, elongation). Compounds were added to such standard reactions as indicated above the gel, i.e. 0.8 μM 2'-FdGTP (lane **5**), 0.8 μM 2'-FdGTP plus 5 μM CTP, ATP, UTP (lane **6**), 0.8 μM GTP (lane **7**), 0.8 μM GTP plus 5 μM CTP, ATP, UTP (lane **8**). For lanes **9** and **10** 0.8 μM 2'-FdGTP or 0.8 μM GTP were preincubated for 30 min with RNP-primer complexes before 5 μM CTP, ATP, UTP were added. Lanes **11-16** correspond to lanes 5-10 with 4 μM instead of 0.8μM 2'-FdGTP and GTP respectively. Lane **17** corresponds to a transcription reaction with 5μCi [α^{32}P]-GTP and 100 μM 2'-FdCTP, lane **18** included 5μCi [α^{32}P]-GTP, 100 μM 2'-FdCTP and 5 μM CTP, ATP, UTP. Lanes **19** and **20** correspond to lanes 17 and 18 with 100 μM CTP replacing 2'-FdCTP.

III.5 Discussion

In vitro transcription assays based on the incorporation of labelled nucleotides into acid insoluble material or on separation of the reaction products on polyacrylamide gels were established to characterize and compare different virus and RNP preparations used for structural studies (chapter I) and to study the impact of 2'-deoxy-2'-fluororibonucleotides on different stages of influenza virus specific transcription. Capped, mRNA-like RNAs were produced by *in vitro* transcription with SP6 or T7 polymerase from plasmid DNA. These mRNA analogs function as substrate for the influenza virus endonuclease, which produces short primer RNAs for transcription initiation. In the presence of such mRNA-like molecules and one labelled nucleotide-triphosphate, primer-plus-one-nucleotide RNAs can be visualized on acrylamide gels. After such a transcription initiation reaction, the nascent RNAs can be elongated by the addition of further nucleotide triphosphates. The fluorinated nucleotide analogs 2'-FdGTP and 2'-FdCTP behaved similarly in such *in vitro* transcription assays. There was no significant change in the product pattern from transcription initiation reactions in the presence or absence of compound. The analysis of transcription initiation was slightly complicated by the fact that the respective effectors to be studied could compete with the labelled nucleotide triphosphate used to visualize inhibitory effects. This means that product bands were of lower intensity in the presence of effectors than in the absence of effectors. To overcome this limitation, transcription assays have been performed in different systems, using two different mRNA molecules and using either labelled CTP or labelled GTP in the reactions. In all cases, the conclusions drawn from the experiments concerning the mechanism of transcription inhibition were the same. We obtained no evidence for transcription inhibition during initiation. Both compounds could be incorporated into nascent transcripts with efficiencies similar to the corresponding NTPs. These results suggest that the formation of the first phosphodiester bond, or previous steps, in particular cap-binding and endonuclease activity are not affected by the fluorinated nucleotide analogs. Inhibition of transcription became only apparent during elongation, when both the amount and the length of the RNA products were significantly decreased. This, and the results of limited elongation reactions suggest as the mechanism of inhibition that the compounds are incorporated into nascent RNAs, which leads to a block of transcription, or at least to a strong reduction of elongation efficiency.

Kinetic studies with influenza virus RNPs showed that 2'-FdGTP competitively inhibits incorporation of GTP into trichloroacetic acid insoluble material. Both mRNA and ApG primed transcription reactions behaved similarly and were half-maximally inhibited by about 100 μM 2'F-dGTP (Tisdale et al., 1995). These observations are consistent with endonuclease and initiation independent inhibition of transcription. Taken together, these results are reminiscient of reports concerning HIV and duck hepatitis B virus specific

reverse transcription in the presence of dideoxynucleotide analogs. Antiviral compounds like AZT-TP (3'-azido-3'-deoxythymidine-triphosphate) or 3TC-TP (2',3'-dideoxy-3'-thiacytidine-triphosphate) are both competitive inhibitors with regards to their homologous dNTPs, as determined by incorporation of labelled nucleotides into acid insoluble DNA products, and they inhibit reverse transcription mainly during elongation by chain termination after being incorporated into nascent DNA (Severini et al., 1995; Goody et al., 1991; Huang et al., 1990; Parker et al., 1991; St.Clair et al., 1987).

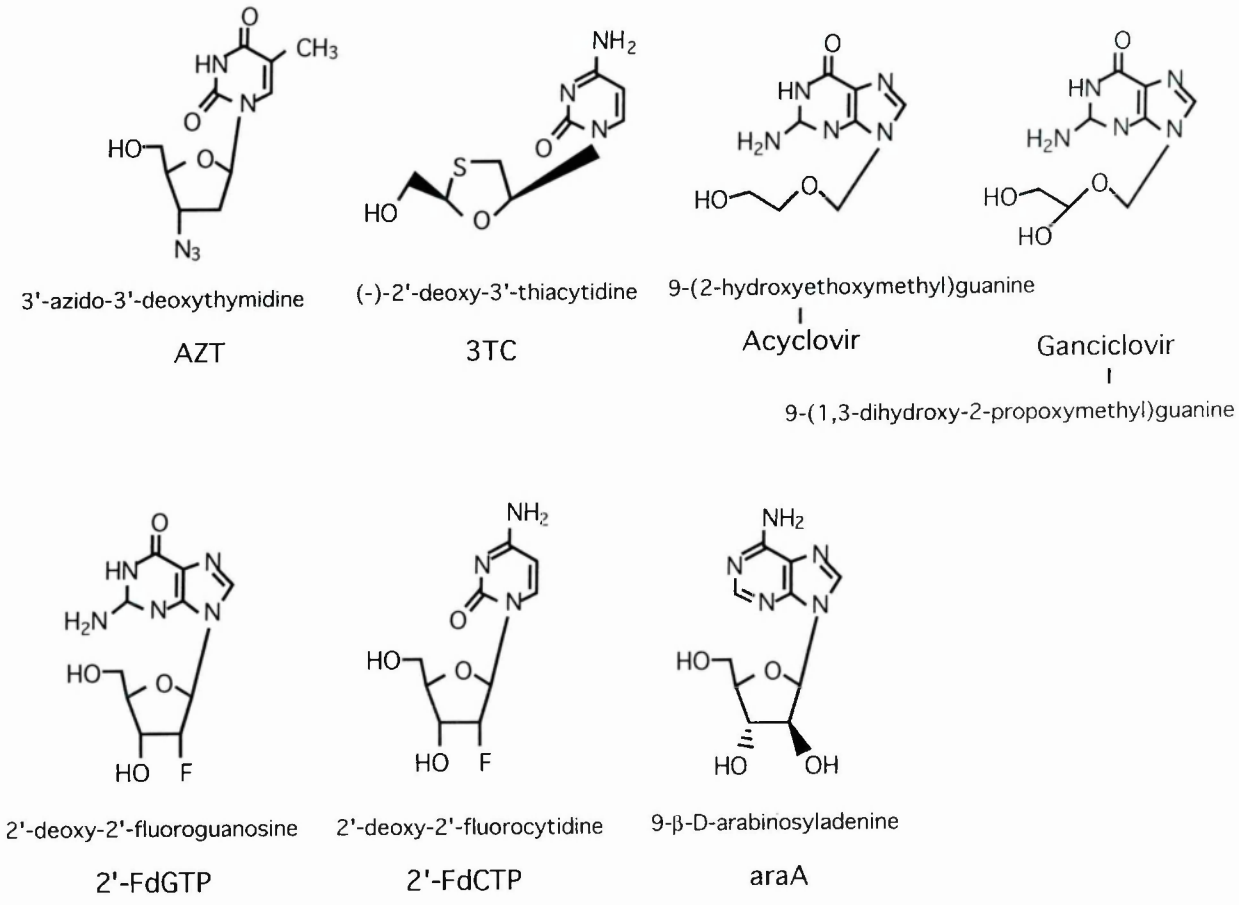


Figure III.14: Structures of antiviral nucleoside analogs. AZT and 3TC are approved from FDA for first-line treatment of HIV infection, Acyclovir, Ganciclovir and araA are active against viruses of the human herpes virus group. Acyclovir is approved for treatment of herpes virus infections, Ganciclovir for cytomegalovirus infections.

Incorporation of dideoxy-nucleotide analogs can lead to suicide inhibition, when the halted polymerase-template-primer complex is stable. This phenomenon is for example observed with inhibition of *E.coli* DNA polymerase by ddTTP or herpes simplex virus DNA

polymerase by the triphosphate of 9-[(2-hydroxyethoxy)-methyl]guanine (acyclovir) (Atkinson et al., 1969; Furman et al., 1984). Studies on the Epstein-Barr virus DNA polymerase, on the other hand, showed considerable reversibility of inhibition by acyclovir triphosphate and in particular, no increase of transcription inhibition by preincubation of template-primer with enzyme and compound, as was seen with *E.coli* and herpes simplex virus DNA polymerases (Datta et al., 1980). The reversibility in the case of Epstein-Barr virus DNA polymerase inhibition is presumably due to a 3'-5' proofreading exonuclease activity of the DNA polymerase. Another example of reversible inhibition is the combination of herpes simplex virus DNA polymerase and 9-β-D-arabinosyladenine triphosphate (araATP) (Derse and Cheng, 1981): araATP does not substitute for dATP during DNA synthesis, but can be incorporated by the polymerase in elongation reactions. Although incorporation of araATP occurs and blocks further elongation, the preincubation of template-primer-polymerase complexes with araATP does not significantly reduce the subsequent incorporation of [^{3}H]dTTP. Derse and coworkers observed maximally 10% reduction of activity after araATP preincubation. The reversibility of the incorporation-block inhibition mechanism in this example is again caused by the 3'-5' exonuclease activity of the herpes simplex virus DNA polymerase. Some of the above described observations are reminiscient of the results from the influenza virus system. Inhibition of influenza virus RNA polymerase by 2'-deoxy-2'-fluororibosides also showed a high degree of reversibility. In [^{3}H]UMP incorporation assays 2'-FdGTP could not substitute for GTP, as expected for chain terminating nucleotide analogs. However, the addition of excess GTP to a transcription reaction resulted in significant UMP incorporation even after extensive preincubation with the compound (Tisdale et al., 1995). The same effect was observed on analytical gels resolving the products of transcription. Preincubation of compound with RNP-primer complexes had no detectable effect on the amounts or sizes of RNA molecules produced by the influenza virus polymerase. In contrast, inhibitory effects during elongation were observed, even with considerable excess of nucleotides in pulse-chase reactions (figure III.10). Is this evidence for a proofreading activity associated with the influenza virus polymerase or are there other possibilities to explain this apparent reversibility of inhibition? In the presence of saturating amounts of primer, transcription inhibition of HIV reverse transcriptase by dideoxy-nucleotides also requires the continous presence of compound. Transcription can resume when large amounts of corresponding dNTP are added to a reaction, and the inhibitory mechanism can therefore resemble reversible inhibition *in vitro,* although it is clear that dideoxy compounds irreversibly block nascent DNAs, and that reverse transcriptase does not possess a proofreading exonuclease activity (Goody et al., 1991). This observation can be explained by the fact that the HIV reverse transcriptase is not trapped in replicative complexes and can continually resume DNA synthesis with new primer molecules. The difference to the influenza virus system is, that the influenza virus RNA-dependent RNA polymerase is apparently able to reuse

previously blocked transcripts. This became for example apparent in the limited elongation reaction, when blocked primer-plus-two RNA products were produced in 30 min preincubation reactions and then partially disappeared in subsequent elongation chase reactions. There are also observations from other laboratories, that support the idea of some kind of proofreading activity to be associated with the influenza virus polymerase. Trimming of primer RNA 3'-ends has been observed *in vitro* after excessive GMP incorporation during transcription initiation. As mentioned above, with increasing amounts of GTP in transcription initiation reactions, the influenza virus polymerase complex can add increasing numbers of G residues to nascent RNAs. After addition of further nucleotides these RNAs are cut down to the right length before being elongated with template specified nucleotides (Ishihama et al., 1986; Plotch et al., 1981).

At present it is not known which subunit of the polymerase is responsible for the observed RNA primer trimming reactions. Although it is possible that the polymerase associated endonuclease might recleave erroneously initiated transcripts back to primer length, Ishihama and coworkers found it unlikely because they observed stringent coupling between removal of excess G residues and transcription elongation, whereas the endonuclease has been shown to be independent of transcription. However, the apparent error rate *in vivo* of the influenza virus polymerase is similar to that of other RNA polymerases lacking proofreading exonuclease functions, so the significance of the apparent primer regeneration activity is at present unclear. The nucleotide analogs may be invaluable tools to study in more detail this unique activity of the influenza virus polymerase, which itself constitutes a promising target of future antiviral compounds. In particular, it will be interesting to determine if the endonuclease activity of the polymerase complex is responsible for the primer RNA regeneration, or if there is yet another active site involved in this process. In the case of fluorinated nucleotide analogs, the apparent primer regeneration reaction can be assumed to lower the antiviral efficiency of the compounds. The successful shut-off of this regeneration mechanism can therefore be expected to increase the inhibitory effect of future polymerase-directed nucleotide analogs.

The above shown experiments could not explicitly establish if the nucleotide analogs were mainly incorporated early in transcription or any time during elongation. This question became important after obtaining several lines of evidence for the polymerase undergoing a conformational change between early and late elongation. It also remained open if incorporation of nucleotide analogs was possible as the first nucleotide, i.e. for the first phosphodiester bond formation. A modified transcription assay will be needed to answer this question. Polymerase preparations that are active on artificial RNA templates and a method to stepwise walk the polymerase along the template will be essential to study transcription elongation *in vitro*. As an alternative, radioactively labelled nucleotide analogs would be useful to analyze in more detail the kinetics of compound incorporation, elongation inhibition and RNA regeneration. They would also help

to answer another question that arises from the above shown results. Whereas dideoxy-nucleotide analogs inevitably cause chain termination after incorporation, antiviral compounds with residual hydroxyl groups can theoretically be internally incorporated into nucleic acid products by polymerase proteins. This has been described with the triphosphate of 9-(1,3-dihydroxy-2-propoxymethyl)guanine (Ganciclovir) and araATP (see above), which are competitive inhibitors of herpes simplex virus and cytomegalovirus DNA polymerases with regard to dGTP. (Derse and Cheng, 1981; Frank et al., 1984; Mar et al., 1985). Due to additional hydroxyl groups these compounds are not absolute chain terminators, but the araAMP-terminated primers are only elongated at a very slow rate, and most primers end up being regenerated by the exonuclease activity of the polymerase. However, internal araAMP residues were detected in viral DNA. For the influenza virus system, labelled nucleotide analogs could help to distinguish between the three possible scenarios of influenza virus specific transcription inhibition by 2'-deoxy-2'-fluororibosides: is inhibition caused by a definitive incorporation block of nascent transcripts, is there no real block, but only a significant decrease of elongation from 2'-FdNMP-terminated primers or is it a combination of both processes, determined by the efficiency of the RNA regeneration activity. Interestingly, these compounds were originally perceived as deoxynucleoside analogs and initially tested for antiviral activity with DNA viruses of the herpes virus group (Wohlrab et al., 1985). Several lines of evidence suggest that also cellular enzymes recognize the compounds as deoxyriboside rather than riboside analogs. Influenza virus inhibition by 2'-FdGTP in cell cultures could only be reversed by the addition of deoxyribosides and deoxycytidine was more effective than cytidine in the reversal of inhibition by 2'-FdCTP (Tisdale et al., 1993). Phosphorylation to triphosphates *in vivo* is essential for antiviral activity of the compounds and it was found that they were good substrates for eucaryotic deoxycytidine kinase *in vitro* (Tisdale et al., 1993). Finally, the nucleoside analogs measurably inhibited DNA polymerase α in vitro (Ki = 48 μM), whereas rat liver RNA polymerase II was not inhibited by up to 850 μM 2'-FdGTP (Tisdale et al., 1995). 2'-FdCTP was found to be inefficiently, but measurably incorporated into DNA molecules by DNA polymerase α, with a K_m of 7 mM compared to a K_m of 0.6 μM for CTP, and more efficiently by AMV reverse transcriptase with a Km of 7 μM compared to 0.14 μM for CTP (Aoyana et al., 1985). Influenza virus specific transcription assays underline the difference between cellular and viral RNA polymerase and demonstrate that 2'-deoxy-2'-fluororibonucleotides can also be recognized as ribonucleotide analogs and be incorporated into RNA molecules.

Materials and Methods

(A) Materials, Cells, Buffers

A1 Viruses and Cells

Influenza virus A/PR/8/34 (H1N1) was purchased from Pasteur-Mérieux, Marcy L'Etoile, France. Influenza virus reassortant strain X31, prepared from infected hen's eggs at Wellcome Research Laboratories, Beckenham, Kent, U.K, was provided by Dr. Martin Ford. X31 virus contains the haemagluttinin and neuraminidase genes from influenza virus A/Aichi/2/68 (H3N2) and the residual genes from A/PR/8/34.

Baculovirus BacPAK6, derived from AcMNPV C6, as well as SF9 (*Spodoptera frugiperda* ovarian cell derived) and High Five™ (*Trichoplusia ni* egg cell homogenate derived) insect cells were provided by Dr. Filippo Volpe and Peter Ertl, Wellcome Research Laboratories, Beckenham, Kent, U.K (Kitts and Possee, 1993). Recombinant baculoviruses containing the coding sequences for influenza virus A/PR/8/34 PB1, PB2 and PA genes were produced as described in chapter II.2.2.

Pichia pastoris yeast strains GS115 (his4), GS115/His$^+$ Mut^S Albumin and GS115/His$^+$ Mut$^+$ β-galactosidase were purchased from Invitrogen, San Diego, Ca.

Escherichia coli bacterial strains:

Strain	source	genotype	reference
BL21 BL21 (DE3)	Novagen	F^- *ompT* [*lon*] *hsdS*$_B$ (r_B- m_B-) DE3, a λ prophage, carries the T7 polymerase gene	Studier, F.W. et al. (1990) Meth.Enzymol. 185: 60-89
HMS174 HMS174 (DE3)	CGSC	F^- *recA*1 *rpoB331 hSDR19* r_{k12}^- m_{k12}^+ Rif^r	Campbell, J. (1978). Proc.Natl.Acad.Sci. USA 75: 2276

Strain	source	genotype	reference
JM109	NEB	F⁻ *traD36 recA1 lacI*^{q} Δ(*lacZ*)*M15* *proA*$^+$*B*$^+$/*rpsL*(Str^r) *thr leu thi lacY galK galT ara fhuA dam dcm supE44* Δ(*lac-proAB*)	Yanisch-Perron, C. et al. (1985) Gene 33: 103-119
XL1-Blue	Stratagene	F'::Tn*10 proA*$^+$*B*$^+$ *lacI*^{q} Δ(*lacZ*)*M15* *recA1 endA1 gyrA96* (Nal^r) *thi hsdR17* (r_k^- m_k^+) *supE44 relA1 lac*	Bullock, W.O. et al. (1987) Bio/Techniques 5: 376-382

A2 Media, growth conditions, transfections and transformations

A.2.1 Sf9 and High Five™ insect cells and baculoviruses

A2.1.1. General cell culture conditions

Insect cells were grown in TC100 medium (GIBCO BRL) supplemented with 10% fetal calf serum (FCS) and 0.1 mg/ml gentamicin (GIBCO BRL). All solutions, plastic flasks, petri dishes, pipettes etc. were of cell culture quality. Culture conditions were as described in (King and Possee, 1992). Sf9 cells were cultured in plastic flasks or in 1l Techne biological stirrer bottles, High Five™ cells only in plastic flasks. For subculturing, the medium from confluent cell monolayers was decanted and the cells dislodged into fresh medium for counting in an improved Neubauer counting chamber (count 25 medium sized squares and multiply by 10^4 to obtain the number of cells/ml). 150 cm^2 flasks were routinely seeded with 2 x 10^7 cells in 10 ml TC100 medium and incubated at 28°C at atmospheric conditions. Subculturing was routinely performed by 1-8 or 1-10 dilution of confluent cell layers. Sf9 cell culture in 1l stirrer bottles is done in 200 ml medium, seeded with 2 x 10^5 cells/ml (about 1 confluent 150 cm^2 flask) and subcultured, when a cell density of 2 x 10^6 cells/ml was reached. Stirrer bottles were maintained at room temperature.

A2.1.2 Freezing and defrosting of insect cells

Cell density was brought to 2 x 10^6 cells/ml with culture medium and cell viability, tested with 0.2% trypan blue staining (cell sample stained with trypan blue before counting; live cells exclude the stain, dead cells are stained), should be high (90-95%). An equal volume of ice-cold freezing mix, i.e. 20% DMSO, 80% growth medium containing 10% FCS, was added to the cells and 1 ml aliquots pipetted into cryogenic vials, immediately placed at -70°C in an insulated freezing box for overnight slow freezing. The cells were then transferred to liquid nitrogen containers. To resusitate cells were rapidly thawed at 37°C and seeded into a 25 cm^2 flask in 10 ml growth medium. After cell attachment (1-2 h), the medium was replaced with fresh growth medium to remove the DMSO.

A2.1.3. Baculovirus propagation

Routine insect cell passaging was done in a room separate from baculovirus propagation to avoid contamination of cell stocks. Virus containing solutions and labware were either decontaminated with 70% ethanol or by autoclaving. Gilson pipettes were regularly decontaminated with detergent and/or 70% ethanol.

A2.1.4 Baculovirus plaque assay to determine virus titre

35 mm dishes (6-well plates) were seeded with 2 x 10^6 cells for 1 h or 1 x 10^6 cells overnight. After cell attachment the growth medium was removed and virus dilutions (e.g. 10^{-3} to 10^{-6}) in 200-500 μl medium was added, with care taken that all cells were well covered with liquid. After 1 h the virus inoculum was aspirated and replaced by 1.5 ml 1% agarose overlay consisting of a mixture of one volume 3% Seaplaque LMP agarose, melted and cooled to 37°C in a water bath, and two volumes of growth medium. The agarose was allowed to set for 15 min and then overlayed with 1 ml growth medium. The plates were placed in a plastic box lined with moist tissue paper and incubated at 28°C for 2 days. At that time a 1:5 dilution of Neutral Red in PBS was added to the liquid overlay and the plates incubated overnight at 28°C. Then the liquid was tipped off and the plates drained in an inverted position for up to 1 h. Virus plaques were picked with a 1 ml Gilson pipette into 500 μl of growth medium and stored at 4°C.

A.2.1.5 Virus stock preparation

Half the volume of a plaque pick was used to infect 2 x 10^6 cells in a 25 cm^2 flask in 4 ml growth medium. After 4 days the virus containing medium was harvested and cleared by centrifugation at 1000 rpm, 10 min, 4°C and stored at 4°C. 0.5 ml of this seed stock was used to infect 5 x 10^6 cells in a 75 cm^2 flask and 10 ml medium resulting in an intermediate virus stock. 0.5 ml of this intermediate virus stock was used to infect 2 x 10^7 cells in a 150 cm^2 flask and 30 ml medium. The maximum virus titre with

recombinant PA expressing viruses was with 2 x 10^6 pfu/ml about 10-fold lower than that obtained with control virus expressing β-galactosidase.

A2.1.6 Infection of insect cells with virus stock for experimental work

For protein expression in flask culture, 3 x 10^7 cells were seeded in a 150 cm^2 flask, the medium removed after cell attachment, and virus inoculum added at a multiplicity of infection (moi) of 10 (10 pfu/cell) in 10 ml medium. After an 1 h incubation at RT, the inoculum was removed and replaced by 30 ml fresh medium. The cells were then incubated at 28°C for 3-5 days before being harvested by centrifugation at 1000 rpm, 10 min, 4°C.

For virus propagation, the same amount of cells was inoculated with 0.1 moi of virus stock to avoid the production of defective interfering particles.

A 2.1.7 Transfection of Sf9 cells by the calcium phosphate method

2 µg BacPAC5 virus DNA was digested with 10U *Sau* I in 100 µl volume at 37°C, 5 h and complete restriction was controlled by comparison with 100 ng circular viral DNA on 0.6% agarose gels. In order to produce recombinant baculoviruses, 1 µg of *Sau* I digested BacPAC5 DNA and 2 µg of transfer plasmid DNA were mixed with 750 µl of transfection buffer (25 mM HEPES, pH 7.1 (±0.05), 150 mM NaCl, 125 mM $CaCl_2$).

The growth medium was removed from 2 x 10^6 cells in a 25 cm^2 flask and replaced by 750 µl fresh medium. The DNA solution was then added dropwise to the cells, and the flask incubated at 28°C for 4 h. The inoculum was removed, the cells washed twice with 5 ml growth medium and incubated at 28°C for 4 days. At the end of this time, the culture supernatant was harvested, clarified by centrifugation and screened in a plaque assay for recombinant viruses. 500 µl of pure virus, a 1:5 and a 1:50 dilution were employed for this purpose on preseeded 6-well plates. Recombinant viruses were characterized by DNA dot-blot assays and Western blots after protein production in Sf9 cells.

A2.2 *Pichia pastoris* cell culture

All cell culture work, transformation and characterization of yeast cells was essentially done as recommended in the Pichia Expression Kit manual of Invitrogen (Catalog no. K1710-01) Yeast cells were usually grown at 30°C and stored at 4°C on RDB agar plates with or without 0.004% histidine. For other purposes, liquid cultures in complex YPD or selective RDB medium were done in standard, baffled culture flasks covered with 2-3 sterile cheese cloth layers. 5 ml medium in 50 ml Falcon tubes were used as standard overnight cultures. Cells from liquid cultures could be harvested and washed with sterile, ice-cold water by centrifugation at 1500 x g, 5 min, 4°C. Methanol utilization (Mut)

phenotypes of cells were determined by comparative growth of colonies on methanol-free MD agar and methanol containing MM agar plates. Colonies were picked from plates with a sterile toothpick and streaked first on MM, then on MD plates.

A2.2.1 Transformation by electroporation

500 ml YPD medium was inoculated with 0.1-0.5 ml of overnight culture and grown for another night to reach an OD_{600} = 1.3-1.5. Cells were harvested, washed twice with 250 ml water, once with 20 ml of ice-cold 1 M sorbitol, before being resuspended in 1 ml 1 M sorbitol. 80 µl of these cells were mixed with 5-10 µg linearized DNA and pulsed in a 0.2 cm cuvette for 10 ms with a field strength of 7500 V/cm, i.e. 1500 V, 50 µF, 200 Ω using an Invitrogen Electroporator II. The cells were rapidly recuperated in 1ml ice-cold 1 M sorbitol and 200 µl spread on RDB plates lacking histidine. Recombinant colonies appeared after 2-4 days incubation at 30°C.

A2.2.2 Protein expression in yeast cells

MUT$^+$: 25 ml BMGY medium in a 250 ml flask was inoculated with a single colony and grown to OD_{600} = 2-6 (16-18 h). The cells were then diluted to OD_{600} = 1 into 100-200 ml BMMY medium in a 1 l flask. Methanol was added to a final concentration of 0.5% every 24 h to maintain induction conditions.

MUT^S: 100 ml BMGY medium in a 1 l flask was inoculated with a single colony and grown to OD_{600} = 2-6 (16-18 h). The cells were then harvested and concentrated to OD_{600} = 10-15 in BMMY medium (approximately 10-20 ml in 100 ml flask). Methanol was continously added as described above.

A2.2.3 Preparation of total yeast cell extracts

(a) glass bead method for small volumes: 100 µl Breaking Buffer and 100 µl acid washed glass beads (SIGMA) were added to the water washed pellet of 1 ml yeast cell culture (centrifugation 1 min bench centrifuge). Cells were lysed by 8 cycles of 30 s vortexing pulses alternating with 30 s recuperation on ice, and cell debris was pelleted during 10 min in a bench centrifuge. Aliquots of the supernatant were mixed with 2x SDS-loading buffer and 2-10 µl used for analytical gel electrophoresis. Breaking Buffer: 50 mM NaH_2PO_4 (6 g/l), 1 mM EDTA (372 mg/l), 5% glycerol (50 ml/l), pH 7.4 set with NaOH.

(b) French Press method: The cells from a 150 ml culture for protein expression were resuspended in 10 ml of the buffer of choice for a following chromatography step, lysed at 10000 psi in a French pressure cell apparatus and cleared by centrifugation at 15000 rpm in SS34 tubes.

A2.2.4 Yeast Media

All media were sterilized by autoclaving at 120°C, 1.2 bar, 20 min.

YPD (rich medium): 1% yeast extract, 2% peptone, 2% glucose, ± 20 g/l agar

RDB (histidine free regeneration medium): 1 M sorbitol, 1% glucose, 1.34% Yeast nitrogen base (YNB), 0.00004% biotin, 0.005% amino acids. [**YNB** (YNB without $(NH_4)_2SO_4$, without amino acids, added together with 100g $(NH_4)_2SO_4$ to 1 l H_2O; **500x biotin stock** solution (0.02%, i.e. 20 mg/100 ml H_2O), stored at 4°C; **100x amino acid stock** solution (0.5%, i.e.500 mg/100 ml H_2 0 of L-glu, L-met, L-lys, L-leu, L-ile)].

MD (histidine free Mut phenotype screen medium): 1.34% YNB, 0.00004% biotin, 1% glucose

MM (histidine free Mut phenotype screen medium): 1.34% YNB, 0.00004% biotin, 0.5% methanol

BMGY/BMMY (protein expression): 1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer, pH 6.0, 1.34% YNB, 0.00004% biotin, 1% glycerol or 0.5% methanol (**1 M potassium phosphate buffer:** 132 ml 1 M K_2HPO_4, 868 ml 1 M KH_2PO_4, pH 6.0 adjusted with phosphoric acid or KOH)

A2.3 *Escherichia coli* cell culture

E. coli cells were routinely grown in LB medium at 30°C or 37°C either in baffled glass flasks, 50 ml Falcon tubes for overnight cultures, or on petri dishes containing LB medium with 15 g/l agar. Cell growth was monitored by absorption measurements at 600 nm (OD_{600}).

A.2.3.1 Transformation of *E.coli* cells

(a) Hanahan method: Bacteria were spread on LB agar plates and single colonies used to inoculate 30 ml SOB medium supplemented with 0.3 ml SOB-complement. Growth was stopped after 2-3 h at 37°C, when the OD_{600} had reached 0.45-0.55. The cells were recovered by centrifugation at 4000 rpm, 10 min, 4°C, washed with 10 ml ice-cold TFB and finally resuspended in 2.4 ml TFB. Then 84 µl DMSO was added and incubated 5 min on ice, followed by the addition of 84 µl 2.25 M DTT, incubated 10 min on ice and 84 µl DMSO added and incubated another 5 min on ice. The cells were then either frozen in liquid nitrogen and stored at -70°C, or used directly for transformation. In the latter case, 210 µl competent cells were incubated with 1-10 µl DNA for 30 min on ice, followed by a 90 s heat shock at 42°C and a 45 min regeneration period in SOC medium at 37°C, before being spread onto LB agar plates containing the appropriate antibiotic, e.g. 100 µg/ml carbenicillin or 50 µg/ml kanamycin.

(b) $CaCl_2$ method: 1 ml of an *E.coli* overnight culture was grown in 100 ml SOB medium until the OD_{600} had reached 0.45-0.55. The cells were incubated on ice for 10 min. Then they were centrifuged at 4000 rpm, 10 min, 4°C, then washed with 10 ml ice-cold 100 mM $CaCl_2$ solution, finally resuspended in 2 ml ice-cold 100 mM $CaCl_2$ and stored 12-24 h at 4°C. 200 µl bacterial suspension was added to 1-10 µl DNA, incubated on ice for 30 min, followed by a 90 s heat shock at 42°C, a 2 min incubation on ice and a 45 min regeneration at 37°C (the last step skipped in the case of ampicillin resistance). The cells were then plated onto prewarmed LB agar plates containing the appropriate antibiotic.

(c) Electroporation

Electroporation of *E.coli* cells was performed on a *E.coli* Pulser™ Transformation Apparatus (BIORAD). A one liter culture of *E.coli* cells at an OD_{600} = 0.5-0.7 was centrifuged in 4 cycles at 4000 x g_{max}, 15 min, 4°C and the pellet gently resuspended in successive steps of 1 l, 0.5 l, 250 ml and 4 ml 10% glycerol to maximally remove salts from the solution. The cells were either stored in 80 µl aliquots at -70°C after rapid freezing in liquid nitrogen, or used immediately. 40 µl cell suspension was mixed with 1-2 µl salt-free DNA, placed in 0.1 cm or 0.2 cm cuvettes and pulsed at 1.8 kV or 2.5 kV respectively. The cells were recuperated in 800 µl SOC medium, incubated for 45 min at 37°C (latter step omitted with apicillin resistance plasmids) and spread on LB agar plates containing the appropriate antibiotic.

A 2.3.2 *E.coli* cell lysis

(a) Sonication: Cell suspensions up to 15 ml were treated with 8 cycles of 15 s pulses at maximal output of a Misonix XL Sonicator and 30 s incubations on ice.

(b) Lysozyme: Cells were harvested in a Sorvall GSA rotor at 7000 rpm, 10 min, 4°C and resuspended in 20 ml Lysozyme buffer (50 mM Tris-HCl, pH 8, 2 mM EDTA, 0.1% TritonX100, 0.2 mM PMSF). Lysozyme was added to 1 mg/ml and incubated at 4°C, 20 min. Then 100 µl 10% DOC, 250 µl 1 M $MgCl_2$ and DNAse I to 0.05 mg/ml were added and incubated at 4°C for 10 min. Cell debris was separated by centrifugation in SS34 tubes at 12000-20000rpm, 15 min, 4°C

A 2.3.3 *E.coli* related Media and Buffers

LB (standard medium): 10 g/l Bactotryptone, 5 g/l Yeast extract 10 g/l NaCl

SOB (transformation): 20 g/l Bactotryptone, 5 g/l Yeast extract, 0.5 g/l NaCl

SOC (transformation): SOB medium plus 250 mM Glucose, 1 mM $MgCl_2$, 1 mM MgSO4, 0.25 mM KCl (200 µl 20% glucose stock solution and 100 µl 100x complement solution (Mg, K salts) were sterile filtrated and added to 10 ml of SOB to produce SOC).

TFB (Hanahan-type transformation): 10 mM Mes-KOH, pH 6.2, 100 mM RbCl, 45 mM $MnCl_2$, 10 mM $CaCl_2$, 3 mM Co(hexamine)Cl_3

TB (protein expression): 12 g/l Bactotryptone, 24 g/l Yeast extract, 4 ml/l glycerol, autoclaved, and supplemented with 1:10 sterile filtrated phosphate buffer (23.1 g/l KH_2PO_4; 125.4 g/l K_2HPO_4)

In certain cases the media were supplemented with 50-100 µg/ml IPTG (in water), 20 µg/ml X-gal (in dimethylformamide), 100 µg/ml carbenicillin or ampicillin (in water), 50 µg/ml kanamycin (in water).

A3 Commonly used buffers

5x ATP-binding buffer	2.5 M KCl 25 mM $MgCl_2$ 5 mM DTT 5 mM EGTA
BC buffer (-)4.5	20 mM sodium citrate, pH 4.5 1 mM EDTA
10x PBS	80 g/l (1.37 M) NaCl 2 g/l (27 mM) KCl 7.7 g/l (43 mM) Na_2HPO_4 1.9 g/l (14 mM) KH_2PO_4
CDI modification buffer 10x (native)	500 mM sodium borate, pH 8-10 100 mM magnesium acetate 900 mM KCl
50x Denhardt's	1% BSA 1% polyvinylpyrrolidon-40 1% Ficoll-400
DMS modification buffer 10x (native)	500 mM sodium cacodylate, pH 7.5 1 M KCl 50 mM $MgCl_2$
5x DNA loading buffer	30% glycerol bromophenol blue
KTT/KTN	100 mM KCl 50 mM Tris-HCl, pH 7.5 1% TritonX100 / 1% NP40
2x Loading buffer proteins	100 mM Tris-HCl, pH 6.8 20% glycerol 4% SDS 1.43 M β-mercaptoethanol small amount of Bromophenol blue
Modification stop buffer 10x	200 mM Tris-HCl, pH 7.5 10 mM EDTA 5% SDS

RNA extraction buffer	0.75 M ammonium acetate 0.1% SDS 0.1 mM EDTA
5x RNA Ligase buffer	250 mM Hepes-NaOH, pH 7.5 50 mM $MgCl_2$ 15 mM DTT
RNA loading buffer, denaturing	20% saccharose 7 M urea
5x RTB	250 mM Tris-HCl, pH 8.3 30 mM $MgCl_2$ 200 mM KCl
20x SSC	3 M NaCl 0.3 M sodium citrate, pH 7
T1 buffer (+)4.5	20 mM sodium citrate, pH 4.5 1 mM EDTA 7 M urea xylencyanol, bromophenol blue
10x TBE, pH 8.3	108 g/l Tris base 55 g/l boric acid 9.3 g/l EDTA
10x TE	100 mM Tris-HCl, pH 8 10 mM EDTA
5x transcription buffer T7/SP6/T3	200 mM Tris-HCl, pH 8 30 mM $MgCl_2$ 50 mM DTT 10 mM spermidine 50 mM NaCl
U2 buffer (+)3.5	20 mM sodium citrate, pH 3.5 1 mM EDTA 7 M urea xylencyanol, bromophenol blue

A4 Plasmids and oligonucleotides

A4.1 Plasmids for protein expression

pXA3 (3.3kb), Promega, Ampicillin resistance, ColE1 ori. Derivatives:

pXA3-PA-wt (5.5kb), contains the complete PA coding sequence (obtained by PCR from pAc373-PA) in the *Sac* I/*Eco* RI sites (II.2.1.2)

pXA3-His-PA (5.6kb) corresponds to pXA3-PA-wt with additional 6 His residues encoded at the C-terminus.

pXA3-NT (3.4kb) contains the PA sequence 1-641, encoding amino acids 1-214 plus an additional peptide (QLGSGTDIRSPGAAAIWFYSVT) at the C-terminus.

pRSETa (3kb), Invitrogen, Ampicillin resistance, ColE1 ori. Derivatives:
pRSETa-PA (5.2kb), contains the complete coding sequence of PA in the *Bam* HI/*Hin* dIII sites (amino acids 1-716).
pRSETa-PA.N (4kb) contains the PA sequence 1-1037, encoding amino acids 1-346 plus an additional peptide (LVPWNSKLDPAANKARKEAELAAATAEQ) at the C-terminus.
pRSETa-PA.L (4.2kb) contains the PA sequence 642-1952, encoding amino acids 215-651 plus an additional peptide (KLDPAANKARKEAELAAATAEQ) at the C-terminus.
pRSETa-PA.E (3.7kb) contains the PA sequence 1176-1952, encoding amino acids 393-651 plus an additional peptide (KLDPAANKARKEAELAAATAEQ) at the C-terminus.
pMALc2 (6.7kb), NEB, Ampicillin resistance, ColE1 ori. Derivatives:
pMAL-PA-M2 (8.8kb), contains the complete coding sequence of PA.
pMAL-PA-NT (6.8kb), contains the PA sequence 1-641, encoding amino acids 1-214, plus an additional peptide at the C-terminus.
pRSETa and pMALc2-derived plasmids coding for PA subunits were provided by Dr. Martin Ford.
pTAg (3.8kb), R&D systems, ampicillin resistance, ColE1 ori.
plasmid used for subcloning of PCR products exploiting the single A overhangs created by non-proofreading DNA plymerases like *Taq* or *Tub* polymerase.
pUBS520 (5.4kb), Boehringer, (Schenk et al., 1995), kanamycin resistance, pACYC ori. The plasmid contains the argU gene, coding for a rare *E.coli* arginine specific tRNA.
pAc373-PA, pAc373-PB1, pAc373-PB2 (Matsuura et al., 1987; St. Angelo et al., 1987); baculovirus transfer vectors containing the complete coding sequences of the influenza virus A/PR/8/34 cloned into the unique *Bam* HI site of pAc373 (pUC8 derivative). Plasmids were provided by Dr. Robert Krug.
pHil-D2 (8.2kb), Invitrogen, Ampicillin resistance, ColE1 ori. Derivative:
pHil-D2-PA-wt (10.4 kb), containing the complete coding sequence of PA (obtained by PCR from pAc373-PA) cloned into the unique *Eco* RI site.

A4.2 Plasmids for *in vitro* RNA synthesis

pGEM7Zf+ (3kb), Promega, ampicillin resistance. Derivatives:
pGEM-PA-wt contains the complete PA coding sequence of PA (obtained by PCR from pAc373-PA) cloned into the *Apa* I/*Eco* RI sites with pGEM-PAL219F being a point mutant of the same plasmid.
pGEM-PANT contains the PA coding sequence from nucleotides 1-653, encoding amino acids 1-218, cloned into the *Apa* I/*Eco* RI sites.

pGEM-PACT contains the PA coding sequence from nucleotides 1377-2148, encoding amino acids 460-716, cloned into the *Apa* I/*Eco* RI sites.

pUCT7GEM (2.7kb) is a derivative of pUC18 (NEB) with the sequence TAAACGACTCACTATAGGGCGATCAAGCTATGCATCTCGAGGAATTC between the *Hin* dIII and *Bam* HI sites. The additional sequence adds a T7 promoter, an influenza virus compatible mRNA primer sequence and *Nsi* I, *Xho* I, *Eco* RI sites to pUC18. The plasmid produces RNA by T7 polymerase transcription to be used in influenza virus endonuclease and transcription assays.

pVN1, (Baudin et al., 1994), contains the sequence for a 81 nt RNA with the influenza virus specifc, conserved vRNA 3' ends ("panhandle RNA") under the control of a T3 polymerase promoter.

A4.3 Oligonucleotides

All chemically synthesized DNA oligonucleotides, that have been used in this study, have been purchased in purified form from either Genosys, Camebridge, U.K. or Eurogentec, Belgium. RNA oligonucleotides have been synthesized and HPLC purified by Viviane Adam at EMBL Heidelberg (sequences shown in II.4, page 74, table II.2). The following table comprises all oligonucleotides related to the present study, even if not all of the experiments in which they were employed appear in the previous chapters. The aim is to give an overview of all the PA-specific oligonucleotides, that are directly available for forthcoming studies.

Primers, specific for the influenza virus A/PR/8/34 PA gene sequence

Name	sequence (5' to 3')	position in the PA sequence	comments
PAGEMAPA1	GAATTCGGGCCCACCATGGAAG ATTTTGTGCGACAATG (38nt)	1-23	EcoRI, ApaI and ACC Kozak sequence at 5' end.
PAXA3SAC5	TCGAGCTCATCGAAGGTCGCATG GAAGATTTTGTGCGACAA (41nt)	1-21	SacI and 12nt pXA3 intermediate sequence at 5' end.
PA235R	GTGCATTAGGATCACCAAGTTC (22nt)	211-190	silent A204T mutation.
504RSFIPA	GTCGGCCTTTGTGCCATT (18nt)	480-462	PA unique BglI site; silent G477T mutation deletes SfiI site.
PAStuI	AGCAGAGGCCTCTGGGAT (18nt)	550-567	PA unique StuI site.
PAXA3HIND3	TGGTCGGCAAGCTTGCGCATT (21nt)	650-630	PA unique HindIII and FspI sites.
PAHINDRS	GAATTCCTAACTTTGGTCGGCA AGCTTGCG (30nt)	654-634	PA unique HindIII site; EcoRI and Stop codon at 5' end.
904FRCAPA	GACCCAAGTCATGAAGGAG (19nt)	898-880	
PACTM460F	GAATTCGGGCCCACCATGAAGG GGGTGTACATCAATAC (38nt)	1378-1400	EcoRI, ApaI and ACC Kozak sequence at 5' end.

PAAflII	AAGGAAGATCCCACTTAAGG (20nt)	1517-1536	PA unique AflII and BstYI sites.
PASA	CATTCCTTAAGTGGGCTCTTCCT TTTATGA (30nt)	1540-1511	PA unique AflII site, BstYI site lost; mutation A1525C, i.e. S509A.
PARK	CATTCCTTAAGTGGGATTTTCCT TTTATGA (30nt)	1540-1511	mutation C1523T, i.e.R508K.
PARG	CATTCCTAAAGTGGGATCCTCCT TTTATGA (30nt)	1540-1511	mutation T1522C, i.e R508G.
1920FPA	CATTGGGAAGGTCTGCAG (18nt)	1896-1913	Pst I site.
1937RPA	CTGCAGACCTTCCCAATG (18nt)	1913-1896	
PACTBGLECOR	AGATCTGAATTCCTAACTCAAT GCATGTGTAAG (33nt)	2131-2148	BglII, EcoRI and Stop sites at 5' end.
PAHISCR	AGATCTGAATTCCTAATGATGA TGATGATGATGACTCAATGCAT GTGTAAGGAAG (55nt)	2127-2148	BglII, EcoRI, Stop sites and 6 His codons at 5' end.

Other DNA oligonucleotides

Name	sequence (5' to 3')	comments
PrimerNS	CATTTATGCAAGCCTTAC (18nt)	Sequence complementary to nucleotides 100-83 of influenza A/PR/8/34 segment 8 vRNA
PB1GEMAPAKPN	GGTACCAAGGAGGGCCCACCAT GTCAATCCGACCTTACTTTTC (43nt)	Sequence complementary to nucleotides 1-23 of influenza A/PR/8/34 PB1 coding sequence, with Kozak sequence, Apa I and Kpn I sites at 5' end.
PB1CTSMAHIND	CCCGGGAAGCTTCTATTTTTGCC GTCTGAGCTC (33nt)	Sequence inversely complementary to nucleotides 2252-2270 of influenza A/PR/8/34 PB1 coding sequence, with Sma I and Hind III sites at 5' end.
PB11275C	CTGTATTAGGCGTCTCCATCC (21nt)	nucleotides 1246-1262 of influenza A/PR/8/34 PB1 coding sequence; PB1 unique BsaHI site.
PB2GEMAPASMA	CCCGGGAAGGAGGGCCCACCAT GGAAAGAATAAAAGAACTAAG (43nt)	nucleotides 1-23 of influenza A/PR/8/34 PB2 coding sequence, with Kozak sequence, Apa I and Sma I sites at 5' end.
PB2CTNOTXBA	GCGGCCGCTCTAGACTAATTGA TGGCCATCCGAAT (35nt)	Sequence inversely complementary to nucleotides 2262-2280 of influenza A/PR/8/34 PB2 coding sequence, with Not I and Xba I sites at 5' end.
PB2544C	GGAGCCAGGATACTAACATCG (21nt)	nucleotides 517-537 of influenza A/PR/8/34 PB2 coding sequence.
NA1125C	GATAGTAAGTTCTCTGTGGAGG C (23nt)	nucleotides 1105-1126 of influenza A/PR/8/34 NA coding sequence.
HAPR81	GAGGATGAACTATTACTGGACC (22nt)	nucleotides 726-747 of influenza A/PR/8/34 HA coding sequence.

(B) Methods for the RNA modification analysis of viral RNPs

B1 Purification of influenza virus RNPs

B1.1 Centrifugation, concentration and washing of virus particles

Virus particles, purified by sucrose gradient centrifugation of infected egg allantoic fluids, were concentrated and washed with PBS by centrifugation 2 h, 50000 x g, 4°C, i.e. 20 krpm with the Beckman SW27 rotor. The following equations were used to calculate centrifugation conditions for other rotor types. The relative centrifugal force (RCF) [g] is related to the speed in revolutions per minute (rpm):

$$RCF = 11.18 \cdot r[cm] \cdot \left(\frac{rpm}{1000}\right)^2 \qquad (1)$$

To adapt centrifugation protocols to different rotors, both Beckman and Kontron provide as a practical parameter the k-factor, which can be used to calculate the time required to pellet particles of known sedimentation coefficients (s). The smaller the k-factor the greater the pelleting efficiency of the rotor. The k-factors are related to the rotor speed and can be calculated using equations (2) and (3):

$$k = \frac{2.53 \cdot 10^{11}\left[\ln\left(\frac{r_{max}}{r_{min}}\right)\right]}{rpm^2} \qquad (2) \qquad k_{actual} = k\left(\frac{rpm_{max}}{rpm_{actual}}\right)^2 \qquad (3)$$

The centrifugation times can then be determined by equation (4):

$$\frac{k_1}{t_1} = \frac{k_2}{t_2} \qquad (4)$$

Accordingly, influenza virus particles were sometimes pelleted 1.5 h, 36 krpm with the Kontron TST41.14 rotor, or 20 min, 45 krpm with the Beckman SW55 rotor. Virus pellets were washed with PBS to maximally remove traces of ribonuclease activity, that has been reported to be a common contaminant in commercially available sucrose used for gradient centrifugations of virus.

B1.2 Influenza virus genomic RNP preparation

Virus particles were lysed by incubation at RT for 10 min with 1% TritonX100, 1 mg/ml lysolecithin, 5 mM MgCl2, 100 mM KCl, 1.5 mM DTT, 5% glycerol and 10 mM Tris-HCl, pH 8. This mixture was centrifuged through a linear 30-60% glycerol gradient in 100 mM NaCl, 50 mM sodium cacodylate (pH 7.5) or Tris-HCl (pH 8), 1 mM DTT. According to the actual volume of virus lysate the centrifugation was done with the SW27 rotor (25 krpm, 16.5 h, 4°C) or with other rotors under corresponding conditions, calculated as outlined above. The glycerol gradients were loaded with virus lysate to 1/10 of the total volume or 1 ml/cm tube diameter, and fractions were collected

from the bottom using a Buchler Auto-Densi-Flow II fraction collector connected to a peristaltic pump. RNP containing fractions were identified by coomassie staining of proteins after 12% SDS-PAGE. Under the above described conditions, RNPs sedimented to regions of 40-50% glycerol and could therefore be stored at -20°C without freezing. Micro-ProDiCon™ ultrafiltration tubes (Spectrum, Breda, NL) were used for buffer changes or to concentrate RNPs. The ultrafiltration device consists of cylindrical cellulose ester dialysis membranes placed in a plastic chamber with an outlet connected to a vacuum pump. The system showed a nonspecific adsorption of 100-300 µg RNP protein after overnight dialysis/50-fold concentration. DOC treated or salt treated RNPs were centrifuged, dialyzed and concentrated as described for complete RNPs.

B2 Conditions for chemical and enzymatic probing of RNA under native conditions

Generally 20-30 µg RNP or 0.1-5 µg RNA were used in modification experiments. The exact conditions to achieve statistical modification were determined empirically for each RNA preparation. The modification reactions were performed in 20-300 µl total volume and terminated by the addition of 1x Modification Stop buffer. In the case of RNPs and enzymatic reactions, the RNAs were purified by phenol extraction after modification. In all cases the RNAs were ethanol precipitated before being further analyzed by reverse transcription or chemical cleavage.

DMS modification: 0-1 µl DMS (Fluka) in 1x DMS buffer, 5-15 min at 37°C.

Kethoxal modification: 3 µl 20 mg/ml kethoxal in 20% ethanol were added to RNA in 1x DMS buffer, 0-1 h at 20°C. The reaction was brought to 50 mM potassium borate, pH 7, before addition of 1x Modification Stop buffer.

RNAse T1 (Pharmacia): 0-1 U in 1x DMS buffer for 15 min 37°C.

B3 Primer extension analysis of modified positions

After enzymatic and chemical modification reactions, the RNA was eventually phenol extracted, ethanol precipitated, washed and resuspended in 3 µl H_2O. A oligodeoxyribonucleotide was labelled at its 5' end by phosphorylation with [γ^{32}P]ATP and T4 PNK and used as a primer for reverse transcription in the following reaction mixture: 1 µl 5x RTB, 1 µl 0.1 M DTT, 0.1 µl RNAsine (Promega), 1.35 µl 2.5 mM dNTP mix, 10^5 cpm primer DNA, 0.1 µl 10 U/µl AMV reverse transcriptase, H_2O to 5 µl total volume. The reactions were incubated 30 min at 37°C, terminated by adding 20 µl Modification Stop buffer and the RNA hydrolyzed after adding 3 µl 3 M KOH by incubation for 3 min at 95 °C and 1 h at 37°C. The DNA was recovered by centrifugation after adding 6 µl 3 M acetic acid, 100 µl 0.3 M sodium acetate, pH 5.2 and 300 µl ethanol. The pellet

was washed with 70% ethanol, resuspended in 3 µl H_2O and 6 µl formamide loading buffer and half of it loaded onto a 12% urea-PAGE gel (C4.2).

B4 Chemical detection of modified positions

To analyze the positions of N3-methylated cytosines, the modified RNA molecules were 3' end labelled by pCp ligation and repurified on urea-PAGE gels. The precipitated RNAs were resuspended in 10 µl 10% hydrazine, 5 µl 1µg/µl *E.coli* tRNA (Boehringer) and incubated 5 min on ice, before being precipitated with 100 µl 0.3 M sodium acetate and 300 µl ethanol. The washed pellets were resuspended in 20 µl aniline, pH 4.5 (stock solution 100 µl aniline, 60 µl acetic acid, 930 µl H_2O) and incubated 15 min at 60°C in the dark. The RNAs are again ethanol precipitated and washed twice with 70% ethanol to completely remove all traces of aniline, then resuspended in H_2O and RNA loading buffer.

B5 Dideoxy-sequencing of RNA

AMV reverse transcriptase (Appligene) was used for RNA sequencing reactions.
3-6.2 µl (about 5 µg) RNA was mixed with 8 µl 5x RTB, 0.8 µl 1 M DTT, 2 µl RNAsine (Promega), 10^5 cpm 5'-labelled DNA primer, 0.6 µl (12 U) reverse transcriptase and H_2O up to 40 µl total volume. 5 µl of this reaction was added to tubes containing 2 µl ddNTP mixtures (e.g. ddATP mix: 1 mM dCTP, dGTP, dTTP, 250 µM ATP, 100 µM ddATP) and incubated for 30 min at 37°C. The reaction was terminated and analyzed as described in B3.

B6 Enzymatic sequencing of RNA

RNAses for enzymatic sequencing of RNA were purchased from Pharmacia, dissolved to 10 U/µl in H_2O and stored at -20°C. All reactions were performed with 10^5 cpm labelled RNA in 3 µl H_2O, in 9 µl total volume and in the presence of 1 µg *E.coli* tRNA as carrier for 15 min at 55 °C.
RNAse T1 (cuts after G residues): 0.05 U in 4 µl T1 buffer (+); RNAse U2 (cuts after A residues): 0.5 U in 4 µl U2 buffer (+); RNAse PhyM (cuts after A and U residues): 0.5 U in 4 µl T1 buffer (+); RNAse BC (cuts after C and U residues): 0.5 U in 4 µl BC buffer (-). The reactions were stopped in liquid nitrogen, followed by storage at -20°C and heating 1 min 95°C just before being loaded onto analytical urea-PAGE gels. Alkaline hydrolysis RNA ladder: 2 x 10^5 cpm of RNA n 6 µl H_2O are incubated with 0.35 µl 0.5 M NaHCO3, pH 9.5 and 2 µl H_2O for 3min at 95°C. Then RNA loading buffer is added before urea-PAGE analysis (C4.2).

(C) Basic methods for nucleic acid analysis, adapted from (Sambrook et al., 1989).

C1 Preparation of plasmid DNA from *E.coli*

C1.1 Small scale (10-20 µg) plasmid preparation

Sol I: 50 mM glucose, 25 mM Tris-HCl, pH 8, 10 mM EDTA

Sol II: 0.2 M NaOH, 1% SDS

Sol III: 60 ml 5 M potassium acetate, 11.5 ml acetic acid, 28.5 ml H_2O

The bacterial cells from 1.5 ml overnight culture suspension were pelleted by centrifugation for 30 s at maximal speed in a bench centrifuge and the supernatant removed by aspiration. The cells were resuspended in 100 µl Sol I and lysed by the addition of 200 µl freshly prepared Sol II and 3 min incubation on ice. Care was taken to avoid surplus mechanical force in order to avoid the fragmentation of chromosomal DNA. Cell fragments, proteins and chromosomal DNA were selectively precipitated by the addition of 150 µl ice-cold Sol III and 3 min incubation on ice. The plasmid DNA was further purified by phenol extraction, i.e. addition of an equal volume of buffer equilibrated phenol, chloroform, isoamylalcohol (25:24:1 v/v/v, pH 6.6, ICN) and phase separation by centrifugation. The plasmid DNA in the top phase was precipitated by the addition of 900 µl ethanol and recovered by centrifugation at maximal speed in the bench centrifuge for 5 min. The pellet was washed with 70% ethanol to remove traces of salt and phenol and the DNA resuspended in 50 µl H_2O. The Promega Magic Minipreps DNA purification system, based on DNA binding columns, was sometimes used as an alternative to phenol extraction.

C1.2 Large-scale (1-2 mg) preparation of plasmid DNA.

The above mentioned Mini-prep procedure was scaled up to larger culture volumes in the following way: Cells from a 500 ml overnight culture were sedimented at 7000 rpm, 5 min, 4°C, resuspended in 6.5 ml Sol I, lysed by addition of 13 ml Sol II (5 min on ice) and cleared with 6.5 ml Sol III (5-10 min on ice). The plasmid contaning supernatant was recovered after centrifugation of cell lysates at 20 000 rpm, 10 min, 4°C in SS34 tubes and the DNA precipitated by the addition of 0.6 volumes isopropanol (2 min, RT). The pellet obtained after centrifugation at 8000 rpm, 5 min, 4°C was dissolved in 1.5-3 ml 50 mM Tris-HCl, pH 8, 1 mM EDTA and first supplemented with 1.12 g/ml CsCl followed by 80 µl/ml ethidium bromide (10 mg/ml stock). This solution was cleared by centrifugation at 8000 rpm, 10 min, RT, transferred to Beckman Quick-Seal™ tubes and the volume adjusted with 80 µg/ml ethidium bromide, 1.12 g/ml CsCl solution. Isopycnic centrifugation was performed with a Beckman 70.1 Ti rotor at 60000 rpm, 24 h, 20 °C and plasmid DNA containing bands collected with a syringe. Ethidium bromide

was removed by 4 cycles of butanol extraction and the DNA precipitated after dilution with 2 volumes of H_2O by the addition of 2 volumes of ethanol. The DNA was resuspended in TE buffer and the concentration determined with the following estimations (Sambrook et al., 1989):
1 A_{260} unit = 50 µg/ml for double-stranded DNA, 40 µg/ml for single-stranded DNA or RNA, 20 µg/ml for oligonucleotides.

C2 Preparation of genomic DNA from *Pichia pastoris*

2 ml of overnight culture suspensions (OD_{600} = 0.7-1) were centrifuged at 3000 rpm, 30 s in the bench centrifuge, the pellet washed with 500 ml H_2O and the cells lysed by adding 200 µl Breaking Buffer (see A2.2.3), 200 µl acid washed glass beads (SIGMA) and 200 µl equilibrated phenol/chloroform/isoamylalcohol, vortexing 30 s, addition of 200 µl H_2O and placing 5 min into a shaking device at high speed. The DNA was precipitated from the aqueous phase by the addition of 1 ml ethanol, centrifuged 3 min at maximal speed in the bench centrifuge, and the pellet resuspended in 20 µl H_2O supplemented with 0.1 µg/ml RNAse A.

C3 DNA dot blot

C3.1 from yeast genomic DNA preparation

5µl genomic DNA from C2 were incubated for 5 min at 65°C in the presence of 75 µl 0.1 M NaOH, cooled on ice, supplemented with 80 µl 20x SSC buffer and filtered onto Hybond N membranes (Amersham), presoaked in 20x SSC, 0.3 M sodium acetate, pH 5.2. The DNA was fixed by baking for 30 min at 120°C.

C3.2 from baculovirus DNA in intact insect cells

A 96-well plate was seeded with 3 x 10^4 cells/well and infected with virus (e.g. 1/10 volume of a virus plaque pick). After infection, the cells were incubated at 28°C for 3-4 days, then washed with PBS, harvested and disrupted by the addition of 40 µl 100 µg/ml proteinase K in TE buffer and 0.1% SDS (37°C, 30 min). The DNA was denatured by the addition of 20 µl 0.5 M NaOH and incubation at 65°C, 5 min, followed by cooling on ice and neutralization with 55 µl 20x SSC, 0.3 M sodium acetate, pH 5.2. Transfer to membrane as in C3.1.

C3.3 Hybridization and Detection

C3.3.1. Radioactively labelled DNA probe:

Prehybridization was performed for 1-4 h, 68°C in 6x SSC, 5x Denhardt's, 0.5% SDS. Radioactively labelled DNA probe (e.g. produced by PCR or random labelling procedures)

was added to an activity of 1-10 x 10^6 cpm/ml and incubated 2-12 h at 68°C. The filters were usually washed twice, 5 min with 2x SSC, 0.1% SDS, twice, 15 min 0.1% SSC, 0.1% SDS before being wrapped into clingfilm and analyzed by autoradiography.

C3.3.2. Non-radioactive labelling and detection with the DIG system:

The DIG system (Boehringer) is based on modified nucleotides carrying the steroid hapten digoxigenin, which can be incorporated into DNA during PCR reactions, thereby generating stable and reusable DNA probes. The hybridization conditions were as described in C3.3.1, except for the use of a specific blocking reagent provided in the kit system. Hybridized, DIG labelled nucleic acids are detected in an enzymatic assay using alkaline phosphatase coupled anti-DIG antibodies with any available alkaline phosphatase substrate, e.g. NBT/X-phosphate. Antibody binding, washing and detection was according to the DIG system manual and similar to the standard development of Western blots (see D4).

C4 Electrophoretic separation and recovery of nucleic acids from gels

C4.1 Agarose gel electrophoresis (AGE)

Submerged, horizontal AGE was performed to separate DNA or RNA molecules of moderate to large sizes (0.2-15 kb). 0.8% w/v agarose gels were routinely prepared by boiling a suspension of nucleic acid grade agarose powder in water, cooling, adding 10x TBE to a final concentration of 1x TBE and pouring the solution into special rectangular gel-formers (e.g. Eurogentec). Gels were run at 50-150 V according to gel size and nucleic acids visualized by submerging the gel in a 1 μg/ml ethidium bromide solution follwed by fluorescence detection in UV_{238nm} through-light. DNA was usually recovered from Seaplaque™ low melting point agarose gel slices after breaking the gel mechanically with a sterile plastic bar and incubation at -80°C or in liquid nitrogen for 1-30 min. The gel slurry was then transferred to a SpinX filtration unit (Costar, Cambridge, MA), the DNA containing solution centrifuged through the filter and the DNA precipitated by the addition of 1/10 volume 3 M sodium acetate, pH 5.2, 2.5 volumes ethanol.

C4.2. Denaturing polyacrylamide gel electrophoresis (urea-PAGE)

Urea-PAGE was used for RNA and DNA sequencing reactions, RNA modification analysis, analysis of oligonucleotides and products of *in vitro* transcription reactions. Premixed 40% acrylamide/bisacrylamide (19:1) and 6-8% Sequagel™ solutions (National Diagnostics, Atlanta, GA) were used to prepare PAGE gels supplemented with 7 M urea. Electrophoresis was performed in 1x TBE buffer at 40 mA. Analytical gels were usually fixed by incubation in 10% ethanol, 10% acetic acid, 20 min, transferred onto Whatman 3MM filter paper and dried at 80°C under vacuum. Nucleic acids were recovered from

urea-PAGE gels by incubation of gel slices overnight at RT in RNA extraction buffer (REB) and ethanol precipitation of the supernatant.

Enzymatic modifications of nucleic acids.

C5 Sequence-specific hydrolysis and dephosphorylation of DNA

Restriction endonucleases were used in the manufacturer's recommended buffers at a concentration of 1-3 U/μg DNA in a volume of 20-50 μl. Enzymatic digestion was carried out at the recommended temperature for 1-2 h. Several enzymes were used together, when they had identical or similar buffer requirements. Otherwise, either an adjustment was made to the buffering system, or the DNA was purified and desalted between digestions. A typical DNA endonuclease reaction contained 10 μl DNA (1-10 μg), 5 μl 10x reaction buffer, 2 μl restriction enzyme (10 U/μl), 33 μl H_2O. Shrimp alkaline phosphatase (0.5-1 U; USB) was added to dephosphorylate DNA 5' ends concomitantly with the endonuclease reaction.

C6 Ligation of DNA molecules

Ligation of DNA molecules was carried out with T4 DNA Ligase (Pharmacia). For plasmid preparations, vector and insert DNAs were mixed in molar ratios of 1:1 to 1:5 and incubated overnight at 12-14°C in a 10 μl reaction comprising x μl vector DNA, y μl insert DNA, 2 μl 5x Ligase Buffer (50 mM Tris-HCl, pH 7.6, 20 mM $MgCl_2$, 2 mM DTT, 10% PEG 8000), 2 μl 10 mM ATP and 6-(x+y) μl H_2O.

C7 Ligation of RNA molecules (pCp labelling)

Ligation of radioactive pCp to the 3' ends of RNAs was carried out with T4 RNA Ligase (Pharmacia). 6 μg RNA (6 μl), 8 μl 5x RNA Ligase buffer with 50 μg/ml BSA, 1 μl 1 mM ATP, 3 μl DMSO, 20 μl 10mCi/ml (3000 Ci/mmol) [^{32}P]pCp, 2 μl T4 RNA Ligase were mixed and incubated overnight at 4°C. The RNAs were then purified by urea-PAGE. vRNAs from 30 μg RNP modification reactions were resuspended in 3.5 μl H_2O after phenol extraction and the reaction contained 3 μl 5x RNA Ligase buffer plus BSA, 1 μl 100 μM ATP, 5 μl [^{32}P]pCp, 1.5 μl DMSO and 1 μl T4 RNA Ligase.

C8 Phosphorylation of DNA and RNA 5' ends

1-5 μg nucleic acid (1-50 pmol) was incubated in a total volume of 10 μl 30 min at 37°C with 1 U T4 polynucleotide kinase (PNK, Pharmacia), 1 μl 10x PNK buffer (one-for all buffer, Pharmacia) and 1-5 μl 3000 Ci/mmol [γ^{32}P]ATP. Labelled nucleic acids were either purified by urea-PAGE or Nuc-Trap™ (Stratagene) gel filtration chromatography.

C9 Dideoxy-sequencing of double-stranded DNA

The sequencing reaction was performed using the Sequenase™ version 2.0 sequencing kit, which includes as the polymerase protein a variant of the bacteriophage T7 DNA polymerase (USB). The sequencing procedure of the kit manual was slightly changed for convenience. 2-4 µg plasmid DNA (e.g. 8 µl Miniprep DNA) was incubated with 3 µl 2 M NaOH 5 min at RT. Then 2 µl 2 pmol/µl primer DNA, 3 µl 3 M sodium acetate, pH 5.2 and 70 µl ethanol were added and, after centrifugation, the precipitated DNA was washed with 70% ethanol and resuspended in 8 µl H_2O. For the primer extension reaction 2 µl Sequenase buffer Rxn, 0.4 µl dGTP, 1.6 µl H_2O, 0.5 µl 600 Ci/mmol [α^{35}S]dATP, 1 µl 100 mM DTT, 0.25 µl Sequenase and 1.75 µl dilution buffer were mixed and added to the DNA for a 5 min incubation at RT. Then 3.5 µl of the reaction were spotted onto a petri dish, prepared with 4 spots of 2.5 µl ddNTPs (and incubated for a further 5 min at 37 °C. The reaction was terminated by the addition of 4 µl Stop solution to each drop, 2 µl were loaded onto a pretemperated 55°C, 6% urea-PAGE gel in a BIORAD sequencing apparatus. After the run the gel was dried for 1 h at 80°C on the glass plate and exposed overnight to Bio-Max X-ray film (Kodak).

C10 Amplification of DNA by the polymerase chain reaction (PCR)

PCR reactions were used in this study to amplify PA coding sequences, to add restriction sites of choice to the ends of PA sequences and to produce PA deletion mutants. PCR reactions consist of cycles of thermal DNA template denaturation, template-primer annealing and primer elongation. Thermostable DNA polymerases are used, that withstand the elevated temperatures during template denaturation. The PCR reactions were performed with a Biometra TRIO thermoblock according to the following standard protocol:

Denaturation of template and primer in reaction mixture at 95°C for 4 min. Addition of 2-5 U *Tub*™ DNA polymerase (Amersham) or *Taq*™ polymerase (Stratagene) and initiation of reaction cycles, i.e. 45 s 95°C denaturation, 1 min 50-68°C annealing, 2 min 72°C primer extension. Between 10 and 25 cycles were run according to the experiment. The annealing temperature was chosen 2-5°C below the theoretical T_m, calculated with the equation T_m (°C) = 4(G+C) + 2(A+T). 100 µl PCR reactions contained PCR buffer (Amersham), 50 mM Tris-HCl pH9, 20 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 250 µM dNTP mix, 50-100 pmol primer DNA, 0.5-2 µg template DNA. The mixture was overlayed with paraffin oil during the PCR reaction. PCR products were analyzed by AGE and ethidium bromide staining. The incorporation of labelled dNTPs during PCR reactions resulted in the production of probes for DNA dot blots.

C11 *In vitro* RNA synthesis with DNA-dependent RNA polymerases

Plasmids containing promoters for the RNA polymerases of bacteriophages T3, T7 or SP6 were linearized with an appropriate restriction enzyme just downstream of the desired 3' end of the RNA coding sequence. Standard transcription reactions contained 40 mM Tris-HCl, pH 8, 20 mM DTT, 15 mM $MgCl_2$, 10 mM NaCl, 1 U/μl RNAsine, 3 mM NTPs, 1% TritonX100, 0.2 μg/μl linearized plasmid DNA, 1-2 U/μl RNA polymerase. To produce capped RNAs, the concentration of GTP was lowered to 0.3 mM and 1.5 mM m7GpppG was included in the reaction. To produce labelled RNAs, the NTP concentration was lowered to 0.5 mM and UTP was replaced with 36 μM UTP and 0.2 μM [α^{32}P]UTP. The RNAs produced in this way were purified by denaturing gel electrophoresis and eluted with RNA extraction buffer. The RNA concentration was determined by measuring the OD_{260} after ethanol precipitation and resuspension in H_2O using the following estimation: 1 OD_{260} = 40 μg/ml RNA

C12 *In vitro* RNA capping reaction

Capped RNAs were either produced by including cap analog in RNA synthesis reactions (see C11, i.e. reduced yield of RNA) or with vaccinia virus capping enzyme in a reaction containing in 15 μl total volume 5-20 μg RNA, 50 mM Tris-HCl, pH 8, 1.25 mM $MgCl_2$, 6 mM KCl, 0.1 μg/μl BSA, 5 mM DTT, 1 U/μl RNAsine, 100 μM S-adenosylmethionine, 1 μM [α^{32}P]GTP, 0.1 U/μl guanylyltransferase. The reaction was incubated for 1 h at 37°C and capped RNA purified by denaturing acrylamide gel electrophoresis.

(D) Basic methods for protein analysis

D1 Determination of protein concentration

D1.1 Absorption of UV light

Protein concentrations were estimated by measuring the absorbance at 280 nm and using the Lambert-Beer law $A = \varepsilon \cdot l \cdot C$, with A being the absorption, ε the molar absorption coefficient ($M^{-1} \cdot cm^{-1}$), l the pathlength (cm) and C the protein concentration (M). The following equation, based on the number of chromophoric amino acids in a specific protein sequence, has been described to well predict the $\varepsilon(280)$ of folded proteins in water (Gill and von Hippel, 1989; Pace et al., 1995):

$$\varepsilon(280)\ [M^{-1} \cdot cm^{-1}] = (\#Trp) \cdot 5500 + (\#Tyr) \cdot 1490 + (\#Cys) \cdot 125$$

D1.2 Coomassie staining

Protein concentrations were estimated by coomassie staining of SDS-PAGE gels and band density comparison to known amounts of protein on the same gel. Alternatively, 200 µl protein dilution was mixed with 800 µl Bradford solution (0.01% w/v Coomassie blue G250, 4.75% v/v ethanol, 8.5% v/v phosphoric acid), absorption measured at 595 nm and compared to a protein standard series, usually made with BSA.

D2 Denaturing protein polyacrylamide gel electrophoresis (SDS-PAGE)

The BIORAD Mini-Protean system or the equivalent Hoefer gel stand were used to cast 7.5-12% discontinous SDS-polyacrylamide gels of 0.75-1.5 mm thickness. National Diagnostics 37.5:1 premixed acrylamide/bisacrylamide solution was used to prepare the separating gel (with 375 mM Tris-HCl, pH 8.8, 0.1% SDS) and the 4% stacking gel (with 125 mM Tris-HCl, pH 6.8, 0.1% SDS) adapted from (Laemmli, 1970)Laemmli, 1970. Protein samples were denatured in reducing SDS-Sample-Buffer and electrophoresis performed at 40 mA in 1x Running Buffer.

D3 Staining of proteins in SDS-PAGE gels

D3.1 Coomassie staining

Acrylamide gels were fixed and stained in 0.25% Coomassie blue R250, 50% ethanol, 10% acetic acid and destained in 7% acetic acid, 5% ethanol.

D3.2 Silver staining

For silver staining the gel was first soaked in 50% ethanol, at least 2 times 1 h, then incubated for 15 min with staining solution (1.6 g $AgNO_3$ in 8 ml H_2O added dropwise to 42 ml 0.36% NaOH, 3.5 ml 28% ammonia, then brought to 200 ml with H_2O), washed thoroughly with H_2O and developed with 0.0025% citric acid, 0.025% formaldehyde. The colouring reaction was stopped by washing with H_2O or destaining solution (D3.1).

D4 Western blot

In Western blotting electrophoretically separated proteins are transferred from a gel to a protein binding membrane. Homemade electrophoresis tanks with gel-membrane sandwich holders or a BIORAD semidry blotting apparatus were used for the electrophoretic protein transfer, which was in both cases performed in Tris-glycine transfer buffer (3 g/l Tris base, 14.4 g/l glycine). Efficient transfer was obtained with 150 V, 3h or 30 V, overnight with wet blotting and 15 V, 30 min semidry blotting. After the transfer, Hybond C nitrocellulose membranes (Amersham) were blocked for 30 min to 1 h in 5% w/v milk, washed with PBS and incubated for 1 h at room temperature with primary antibody. Primary antibody was removed in PBS and PBS/0.5% NP40 washing steps before peroxidase-coupled secondary antibody or radioactively labelled protein A were added diluted 1:1000 in 5% milk for 1 h at room temperature. The blot was developed after washing with PBS and PBS/0.5% NP40 either with the chemoluminescence kit of Amersham or by incubation with 10 mg diaminobenzidene, 0.01% H_2O_2.

D5 Protein concentration and dialysis

D5.1 Ammonium sulfate precipitation

Ammonium sulfate was used at saturation (697 g/l) to quantitatively precipitate proteins from solution, or at lower concetrations for limited fractionations, according to standard tables, e.g. (Englard and Seifter, 1990). The desired amount of solid ammonium sulfate was added to protein solutions, which were then incubated at 4°C for at least 1 h, before being centrifuged and resolubilized.

D5.2 Dialysis

Spectra/Por® dialysis membranes (Spectrum, Breda, NL) were pretreated by boiling for 10 min in 2% $NaHCO_3$, 1 mM EDTA, then thoroughly washed with H_2O and stored in 20% ethanol, 0.05% NaN_3 at 4°C.

D5.3 Ultrafiltration

Micro-ProDiCon™ tubes (see B2.1) were used for simultaneous dialysis and concentration of protein solutions up to 50 ml, Centricon tubes (Amicon) were alternatively used at 4000 x g for the concentration of solutions up to 5 ml.

D6 Circular Dichroism Spectroscopy

CD spectra were collected on a JASCO 600 spectrapolarimeter in a 1 mm pathlength rectangular cuvette thermostatted at 25°C in 50 mM sodium phosphate, pH 7.6, 50 mM NaCl, 2 mM DTT. The measurement was performed at a scan rate of 50 nm/min, 1 s integration time and 2 nm band width.

D7 Chromatographic methods

D7.1 Nickel chelate affinity chromatography

The purification of histidine-tagged polypeptides was done according to the recommendations in the Novagen pET system manual. Affinity columns were prepared with Chelating Sepharose 6B (Pharmacia) and 15 ml plastic columns (Pierce), charged with 50 mM $NiSO_4$. Chromatography was performed in 5 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl, pH 7.9 (native conditions) or in the presence of 8 M urea (denaturing conditions). Proteins were eluted with a 30-100 mM imidazole gradient.

D7.2 Streptavidin affinity chromatography

The purification of biotin-tagged polypeptides was done according to the recommendations in the Promega PinPoint system manual. 1 ml SoftLink resin columns (1 cm diameter) were prepared and run in 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10% glycerol. Bound proteins were eluted in the presence of 5 mM biotin.

D7.3 Amylose affinity chromatography

Amylose resin was purchased from NEB and chromatography on 2 ml columns performed in 50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM EDTA, 10 mM β-mercaptoethanol, 5% glycerol. Columns were usually washed with 1 M NaCl and bound protein eluted with 20 mM maltose in the same buffer.

D7.4 Ion exchange chromatography

Ion exchange chromatography was carried out using MonoQ or MonoS resins in prepacked columns from Pharmacia (HR 5/5) or BIORAD (Econo-PAC cartridges) at a flow rate of 1 ml/min. The columns were run with the following buffers: A, 0-100 mM NaCl, 5 mM MgCl2, 1 mM PMSF, 1 mM DTT, 50 mM Tris-HCl, pH 7.5 or 20 mM Bis-Tris-HCl, pH 6; B, buffer A plus 1-1.5 M NaCl).

(E) Protein activity assays

E1 Influenza virus RNP transcription assay (see III.2)

E1.1 TCA precipitation assay

Approximately 20 µg of detergent disrupted virus, or 0.03-3 µg purified RNPs were added to a reaction mixture containing, in a final volume of 50 µl: 50 mM Tris-HCl, pH 7.8, 100 mM KCl, 10 mM NaCl, 1 mM DTT, 5 mM MgCl2, 0.4 mM ApG or 100 ng primer RNA, 0.5 mM GTP, CTP, 1 mM ATP, 50 µM UTP, 0.1 µM 3000 Ci/mmol [α^{32}P]UTP. Reaction mixtures were incubated for 1 h at 31°C, then stopped by the addition of 50 µl 20% trichloroacetic acid (TCA), 2% casamino acids, incubated 30 min on ice, filtered through Whatman GS/A filter disks, washed with 20 ml 10% TCA, dried and quantified by liquid scintillation counting.

E1.2 Gel analysis of transcription products

Reactions similar to the one described in E1.1 were either phenol extracted and ethanol precipitated, or directly denatured in RNA loading buffer, 2 min at 95°C, before being analyzed on 20% urea-PAGE gels.

E2 *In vitro* ATP binding assay, adapted from (Clertant and Cuzin, 1982) .

[α^{32}P]ATP (80 µCi) was oxidized by incubation with 4 mM HCl and 4 mM $NaIO_4$ in 12 µl total volume for 30 min at room temperature in the dark. Then 3 µl 50% glycerol were added to quench the reaction, followed by a further incubation for 20 min, room temperature, in the dark. Finally the reaction was diluted by adding 5 µl H_2O just and directly used for labelling reactions. The labelling reaction was performed for 1-15 min at room temperature in 50 µl volume in the presence of 1x ATP-binding buffer and 0.25 µM oxidized ATP (2 µl). Sometimes unlabelled NTPs or casamino acids were added at this stage. The protein-ATP complexes were stabilized by incubation overnight on ice with 19 mM NaCNBH3 (adding 2 µl 0.5 M stock solution to the binding reaction). The reactions were stopped by the addition of SDS sample buffer and the proteins were analyzed by SDS-PAGE and autoradiography. To analyze ATP binding of the influenza virus polymerase complex, 2 µl oxidized ATP was added to standard transcription reactions (E.1) with different combinations of NTPs.

E.3 *In vitro* ATPase assay

E.3.1 Thin Layer Chromatography (TLC) separation of reaction products.

ATPase activity in protein samples was analyzed by the release of ^{32}P from [γ^{32}P]ATP, or [^{32}P]ADP from [α^{32}P]ATP. Reactions were in 20 μl total volume and contained 50 mM Tris-HCl, pH7.5, 1-10 mM $MgCl_2$, 1 mM DTT and 0.5 μM labelled ATP. After incubation for 1 h at 37°C, the reaction was stopped by the addition of 2 μl 0.5 M EDTA. 1 μl was spotted onto a polyethylenimine (PEI) coated TLC plate, dried, and the nucleotides separated by ascending chromatography in 1 M LiCl, 0.5 M formic acid. The plates were then dried and subjected to autoradiography with Kodak XAR-5 films.

E.3.2 Colorimetric assay

In this assay an ATPase activity is quantified by measuring the rate of phosphate release into solution. The amount of phosphate was determined as a phosphomolybdate-malachite green complex. The ATPase reaction was performed in a total volume of 50 μl with a freshly prepared buffer mix containing 50 mM Tris-HCl, pH 7.5, 1-10 mM $MgCl_2$, and 1 mM DTT. The reaction was started by adding ATP to a final concentration of 1 mM. After an incubation at 37°C between 30 min and 2 h, the reaction was terminated by the addition of 200 μl malachite green reagent, freshly prepared by mixing the stock solutions 5.72% w/v ammonium molybdate in 6 M HCl, 2.32% w/v polyvinylalkohol, 0.0812% w/v malachite green and H2O at a ratio of 1:1:2:2. The coloured phosphate complex was measured after a 5 min stabilization period in an ELISA plate reader using dual wavelength settings of 620 nm and 450 nm for test and reference filters respectively. Inorganic phosphate standards were prepared using 10-200 μM KH_2PO_4 solutions in ATPase buffer mix.

Conclusions and Perspectives

The goal of this study was to provide data for a better understanding of the relationship between the influenza virus RNP structure and the RNP transcriptional activity. The RNPs can be regarded as small RNA synthesis machines, which produce two different kinds of RNA in a regulated fashion by using only a small number of proteins. Although influenza virus RNPs have already been studied for many years, the important determinants of RNP assembly and the mechanisms of regulated RNA synthesis have largely remained elusive. The major problem has been the lack of suitable systems to express and purify single components of the RNP, so that functional reconstitution of the particle has not yet been achieved. In the present study the problem of RNP structure and function has been approached from two different sides. On the one hand, *in vivo* assembled, viral RNPs have been purified from virion preparations and used for structural and functional studies, and on the other hand it was attempted to establish an expression and purification scheme for the PA polymerase subunit in order to obtain evidence for a possible function of this essential, viral polymerase cofactor and to get a step closer to the production and purification of functional, recombinant polymerase protein.

The secondary structure of the influenza virus genomic RNA was studied in the context of the *in vivo* assembled, functional RNP particle using RNA modification analysis. The results suggest that the RNA molecule is held in a single stranded form by nucleoprotein, which binds to the RNA without causing base protection at Watson-Crick positions and prevents the formation of RNA secondary structure elements. The polymerase complex appears to be binding to both conserved vRNA ends of the genome and causes protection of Watson-Crick positions from chemical modification and enzymatic hydrolysis of 15 bases at the 3'-end and at least 14 bases at the 5'-end. In the absence of the polymerase proteins the bases at the vRNA extremities become reactive to modification, suggesting that the vRNA is completely single-stranded on polymerase-free RNPs despite a partial, inverted complementarity of the vRNA ends. The RNPs have been found to partially unwind under both high and low salt conditions and could then be visualized as circular structures by electron microscopy (courtesy of Dr. Rob Ruigrok). In the absence of the polymerase only linear structures are observed. Together, the results suggest that the vRNA ends are brought together by the polymerase to form a ternary complex, but that a double-stranded RNA secondary structure does not exist in the absence of the polymerase. These observations make it unlikely that the formation of a stable double-stranded panhandle structure is needed for polymerase binding or for transcription termination and they will demand a reexamination of the transcription termination mechanism in the future. We have proposed a possible model for an alternative transcriptional mechanism independent of a stable RNA secondary structure, that is also consistent with most of the previously published observations. In this model

the polymerase is expected to stay bound to the conserved vRNA 5' end during transcription, while advancing along the template RNA from the 3' end. To test this model it will be necessary to harvest RNPs at different times after transcription initiation and monitor the accessibility of the bases at the vRNA 5'-end for chemical modification. In parallel, the elongating RNPs may be studied by electron microscopy. Unwinding in high or low salt will then be expected to produce lariat-shaped elongation complexes, provided that the polymerase-template interaction is stable enough under changing ionic conditions. Further evidence for or against the new model will come from topological studies of the RNP, because the proposed mechanism predicts a relative underwinding of the RNP supercoil to drive polymerase elongation, analogous to transcription from supercoiled plasmids. First evidence for the relative underwinding is provided by the fact that the compact RNP structure opens up when ionic strength is changed in either direction, i.e. the RNP supercoil appears to be poised to be unwound. If this is the case, polymerase-free RNPs, that are allowed to freely rotate and unwind beyond the circular stage, will have the opposite type of supercoil than complete RNPs. In addition, this mechanism predicts a novel, topoisomerase-like activity to be required for the formation of transcriptionally active RNPs.

RNP activity assays confirmed that the RNP preparations used for structural studies were functional and in a physiologically relevant conformation. Furthermore, several lines of evidence suggested that the influenza virus polymerase undergoes a conformational change between transcription initiation and elongation. The RNPs produced significant amounts of short transcripts (primer +1nt, primer +2nt) even in complete transcription reactions, presumably from abortive cycling of the initiation complex, as has also been described for other RNA polymerases. The RNPs showed a ten-fold higher K_m for ATP than for the other NTPs and were found to copurify on glycerol gradients with a significant ATPase activity, that was higher in initiation reactions than in limited elongation reactions or complete transcription reactions. Finally, transcription assays in the presence of ATP analogs gave results consistent with the idea that ATP β-γ-bond hydrolysis is required for transcription initiation, but not during elongation reactions. The presumed energy consuming reaction can be expected to be a rate-limiting step in influenza virus transcription and it will be interesting to identify the nature of this reaction in detail, because cap-binding, endonuclease activity and incorporation of the first nucleotide appear to be independent of ATP hydrolysis. For further studies it will be crucial to succeed in the separate preparation of initiation and elongation complexes for comparative studies of these two phases.

RNP activity assays were also used to study the mechanism of influenza virus inhibition by fluorinated nucleotide analogs. The results suggest that both cytidine and guanosine derivatives of 2'-deoxy-2'-fluororibonucleotide triphosphates were

incorporated into nascent mRNAs by the influenza virus polymerase and blocked, or strongly slowed down, elongation from such transcripts. Evidence obtained from pulse-chase reactions and RNP-primer preincubations with the nucleotide analogs suggested that transcription inhibition is not irreversible and implies that influenza virus RNPs possess the ability to regenerate blocked primers. A proofreading-like activity would be unique for a RNA polymerase, and it will be interesting to further determine, if the endonuclease activity of the polymerase is responsible for the apparent reversibility of inhibition. Both an endonuclease assay with labeled primer RNA as well as labeled nucleotide analogs may be useful in the future to approach this question.

The influenza virus RNA polymerase complex is a multifunctional protein. The currently known functions, cap-binding, endonuclease activity and RNA synthesis, are essential for virus multiplication and constitute potential targets for antiviral therapies, but it is likely that a number of other polymerase functions are still to be identified. The active site performing RNA synthesis has been mapped to the PB1 polymerase subunit, PB2 is the cap-binding subunit and presumably harbours the endonuclease activity. However, no function whatsoever has been proposed yet to explain why PA is essential for viral replication. It is not known, which polymerase subunits interact with the genomic RNA on RNPs and it appears that at least one additional, energy consuming function may be needed to generate the final RNP topology of functional RNPs. In addition, there is evidence for an ATPase activity copurifying with the RNPs, which may be required for abortive cycling during transcription initiation and/or the transition from initiation to elongation. In an attempt to address some of these open questions and to reveal a function for the PA polymerase subunit, several protein expression systems were tested for the production of recombinant PA protein. In most expression systems the PA protein production was very low and there was evidence for significant toxicity in both procaryotic and eucaryotic cells. Extensive fragmentation and insolubility were additional problems encountered in bacterial systems. Both baculovirus infected insect cells and recombinant *Pichia pastoris* clones produced full-length PA protein, but the yields were too low to achieve purification to homogeneity. Taken together, none of the systems appears as yet suited for the expression and purification of full-length PA for structural and functional studies. Affinity purified PA protein from bacterial cells as well as semipurified PA from eucaryotic cells copurified with ATPase activity, which was absent in corresponding fractions of PA-free cell extracts. Because of extensive PA protein fragmentation in *E. coli* and the large amount of residual host cell proteins in PA preparations from eucaryotic cells, it was not possible to exclude that the observed ATPase activity was due to a contaminating protein. Preliminary results with the coexpression of a rare tRNA in *E.coli* suggested that the PA codon bias, which largely differs from that of highly expressed genes in *E.coli*, may be one reason for the low protein yields in the bacterial

system. It may be tried in the future to obtain further evidence for an ATPase activity of PA by the comparison of wild-type PA with parallel preparations of PA proteins containing point mutations in the presumed ATP-binding motif. Eucaryotic systems produced PA protein mainly in full-length form albeit at low levels. It has not been tried to affinity purify PA from these systems with either anti-PA antibodies or after the addition of specific purification tags.

PA protein was also produced by *in vitro* translation in wheat germ extracts. This system was established to determine if PA has RNA binding activity. Preliminary results have been obtained, which suggest sequence specific binding of PA to the conserved 5' end of the genomic RNA and to the conserved sequence near the 5' end of influenza virus specific mRNAs, but not to the 3' end of genomic RNA. Other unspecific RNAs did not bind to PA in this assay consistent with a considerable sequence specificity of the interaction. In addition, molar excess of poly A RNA, poly U RNA or tRNA did not compete with PA binding to the specific RNA sequence. These results support the idea of PA playing a role both as an integral part of the polymerase complex as well as in isolated form. The binding of PA to a conserved, influenza virus specific sequence, which is assumed to be important in transcription and replication of the genome as well as in influenza virus specific translation, gives a first idea of how one protein might influence both viral replication and gene expression. The *in vitro* translation system may become useful for a project to determine the exact RNA sequence specificity and to determine important regions on the PA primary sequence involved in RNA binding.

References

(A)

Akkina, R. K., T. M. Chambers, D. R. Londo, and D. P. Nayak. (**1987**) Intracellular localization of the viral polymerase proteins in cells infected with influenza virus and cells expressing PB1 protein from cloned cDNA. *J Virol* 61 (7): 2217-2224.

Akoto-Amanfu, E., N. Sivasubramanian, and D.P. Nayak. (**1987**) Primary structure of the polymerase acidic (PA) gene of an Influenza B virus (B/SING/222/79). *Virology* 159: 147-153.

Andino, R., G. E. Rieckhof, P. L. Achacoso, and D. Baltimore. (**1993**) Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'- end of viral RNA. *Embo J.* 12 (9): 3587-98.

Aoyana, H., L. Sarih-Cottin, L. Tarrago-Litvak, S. Litvak, and W. Guschlbauer. (**1985**) 2'-fluoro-2'-deoxycytidine triphosphate as a substrate for RNA and DNA-dependent DNA polymerases. *Biochim Biophys Acta* 824: 218-224.

Ashley, Ct, Jr., K. D. Wilkinson, D. Reines, and S. T. Warren. (**1993**) FMR1 protein: conserved RNP family domains and selective RNA binding. *Science* 262 (5133): 563-566.

Atkinson, M.R., M.P. Deutscher, A. Kornberg, R. Russel, and J.G. Moffatt. (**1969**) Enzymatic synthesis of deoxyribonucleic acid. XXXIV. Termination of chain growth by a 2',3'-dideoxyribonucleotide. *Biochemistry* 8: 4897-4904.

(B)

Barcena, J., M. Ochoa, S. de la Luna, J. A. Melero, A. Nieto, J. Ortin, and A. Portela. (**1994**) Monoclonal antibodies against influenza virus PB2 and NP polypeptides interfere with the initiation step of viral mRNA synthesis in vitro. *J Virol* 68 (11): 6900-6909.

Barrera, I., D. Schuppli, J. M. Sogo, and H. Weber. (**1993**) Different mechanisms of recognition of bacteriophage Q beta plus and minus strand RNAs by Q beta replicase. *J Mol Biol* 232 (2): 512-521.

Baudin, F., C. Bach, S. Cusack, and R. W. Ruigrok. (**1994**) Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *Embo J* 13 (13): 3158-3165.

Beaton, A. R. and R. M. Krug. (**1981**) Selected host cell capped RNA fragments prime influenza viral RNA transcription in vivo. *Nucleic Acids Res* 9 (17): 4423-4436.

Beaton, A. R. and R. M. Krug. (**1984**) Synthesis of the templates for influenza virion RNA replication in vitro. *Proc Natl Acad Sci U S A* 81 (15): 4682-4686.

Beaton, A. R. and R. M. Krug. (**1986**) Transcription antitermination during influenza viral template RNA synthesis requires the nucleocapsid protein and the absence of a 5' capped end. *Proc Natl Acad Sci U S A* 83 (17): 6282-6286.

Biswas, S. K. and D. P. Nayak. (**1994**) Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. *J Virol* 68 (3): 1819-1826.

Blaas, D., E. Patzelt, and E. Kuechler. (**1982**) Identification of the cap binding protein of influenza virus. *Nucleic Acids Res* 10 (15): 4803-4812.

Blok, V., C. Cianci, K.W. Tibbles, S.C. Inglis, M. Krystal, and P. Digard. (**1996**) Inhibition of the Influenza virus RNA-dependent RNA polymerase by antisera directed against the carboxy-terminal region of the PB2 subunit. *J Gen Virol* 77: 1025-1033.

Bouloy, M., S.J. Plotch, and R.M. Krug. (**1978**) Globin mRNAs are primers for the transcription of influenza viral RNA in vitro. *Proc. Natl. Acad. Sci. USA* 75: 4886-4890.

Bouloy, M., M. A. Morgan, A. J. Shatkin, and R. M. Krug. (**1979**) Cap and internal nucleotides of reovirus mRNA primers are incorporated into influenza viral complementary RNA during transcription in vitro. *J Virol* 32 (3): 895-904.

Bouloy, M., S. J. Plotch, and R. M. Krug. (**1980**) Both the 7-methyl and the 2'-O-methyl groups in the cap of mRNA strongly influence its ability to act as primer for influenza virus RNA transcription. *Proc Natl Acad Sci U S A* 77 (7): 3952-3956.

Braam, J., I. Ulmanen, and R. M. Krug. (**1983**) Molecular model of a eucaryotic transcription complex: functions and movements of influenza P proteins during capped RNA-primed transcription. *Cell* 34 (2): 609-618.

Braam-Markson, J., C. Jaudon, and R. M. Krug. (1985) Expression of a functional influenza viral cap-recognizing protein by using a bovine papilloma virus vector. *Proc Natl Acad Sci U S A* 82 (13): 4326-4330.

Brenner, S., F. Jacob, and M. Meselson. (1961) An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* 190: 576-581.

Bukrinskaya, A.G., N.K. Vorkunova, and N.L. Pushkarskaya. (1982) Influenza virus uncoating in infected cells and effect of rimantadine. *J. Gen. Virol.* 60: 49-59.

Burd, C.G. and G. Dreyfuss. (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science* 265: 615-621.

(C)

Caton, A. J. and J. S. Robertson. (1980) Structure of the host-derived sequences present at the 5' ends of influenza virus mRNA. *Nucleic Acids Res* 8 (12): 2591-2603.

Cavener, D.R. and C.R. Stuart. (1991) Eucaryotic Start and Stop translation sites. *Nucleic Acids Res* 19: 3185-3192.

Chamberlin, M.J. and T. Ryan. (1982) Bacteriophage DNA-dependent RNA polymerases. *Enzymes* 15: 87-108.

Chattopadhyay, D. and A.K. Banerjee. (1987) Phosphorylation within a specific domain of the phosphoprotein of vesicular stomatitis virus regulates transcription in vitro. *Cell* 49: 407-414.

Cho, M.W., O.C. Richards, T.M. Dimitrieva, V. Agol and E. Ehrenfeld. (1993) RNA duplex unwinding activity of poliovirus RNA-dependent RNA polymerase 3D pol. *J Virol* 67: 3010-3018.

Choppin, P.W., J.S. Murphy, and I. Tamm. (1960) Studies of two kinds of virus particles which comprise influenza A2 virus strains. III. Morphological characteristics: independence of morphological and functional traits. *J Exp Med* 112: 945-952.

Chou, J. and B. Roizman. (1992) The gamma 1(34.5) gene of herpes simplex virus 1 precludes neuroblastoma cells from triggering total shutoff of protein synthesis characteristic of programed cell death in neuronal cells. *Proc Natl Acad Sci U S A* 89 (8): 3266-3270.

Chou, P.Y. and G.D. Fasman. (1978) Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* 47: 45-148.

Chu, C.M., I.M. Dawson, and W.J. Elford. (1949) Filamentous forms associated with newly isolated influenza virus. *Lancet* i: 602-603.

Chung, T. D., C. Cianci, M. Hagen, B. Terry, J. T. Matthews, M. Krystal, and R. J. Colonno. (1994) Biochemical studies on capped RNA primers identify a class of oligonucleotide inhibitors of the influenza virus RNA polymerase. *Proc Natl Acad Sci U S A* 91 (6): 2372-2376.

Ciampor, F., P.M. Bayley, M.V. Nermut, E.M.A. Hirst, R.J. Surgrue, and A.J. Hay. (1992) Evidence that the amantadine-induced, M2-mediated conversion of influenza A virus hemagglutinin to the low pH conformation occurs in an acidic trans golgi compartment. *Virology* 188: 14-24.

Cianci, C., L. Tiley, and M. Krystal. (1995) Differential activation of the influenza virus polymerase via template RNA binding. *J Virol* 69 (7): 3995-3999.

Clem, R. J., M. Fechheimer, and L. K. Miller. (1991) Prevention of apoptosis by a baculovirus gene during infection of insect cells. *Science* 254 (5036): 1388-1390.

Clertant, P. and F. Cuzin. (1982) Covalent affinity labeling by periodate-oxidized [α-^{32}P]ATP of the Large-T proteins of polyoma and SV 40 viruses. *J Biol Chem* 257 (11): 6300-6305.

Colman, P.M. (1994) Structure-based drug design. *Curr. Op. Struc. Biol.* 4: 868-874.

Compans, R. W., H. D. Klenk, L. A. Caliguiri, and P. W. Choppin. (1970) Influenza virus proteins. I. Analysis of polypeptides of the virion and identification of spike glycoproteins. *Virology* 42 (4): 880-889.

Compans, R. W., J. Content, and P. H. Duesberg. (1972) Structure of the ribonucleoprotein of influenza virus. *J Virol* 10 (4): 795-800.

Cusack, S., R.W.H. Ruigrok, P.C.J. Krygsman, and J.E. Mellema. (1985) Structure and composition of influenza virus. A small angle neutron scattering study. *J Mol Biol* 186: 565-582.

(D)

Datta, A.K., B.M. Colby, J.E. Shaw, and J.S. Pagano. (1980) Acyclovir inhibition of Epstein-Barr virus replication. *Proc Natl Acad Sci USA* 77 (9): 5163-5166.

de la Luna, S., J. Martin, A. Portela, and J. Ortin. (1993) Influenza virus naked RNA can be expressed upon transfection into cells co-expressing the three subunits of the polymerase and the nucleoprotein from simian virus 40 recombinant viruses. *J Gen Virol* 74 (Pt 3): 535-539.

Derse, D. and Y.-C. Cheng. (1981) Herpes simplex virus type 1 DNA polymerase. *J Biol Chem* 256 (16): 8525-8530.

Desselberger, U., V. R. Racaniello, J. J. Zazra, and P. Palese. (1980) The 3' and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. *Gene* 8 (3): 315-328.

Detjen, B. M., Angelo C. St, M. G. Katze, and R. M. Krug. (1987) The three influenza virus polymerase (P) proteins not associated with viral nucleocapsids in the infected cell are in the form of a complex. *J Virol* 61 (1): 16-22.

Dhar, R., R. M. Chanock, and C. J. Lai. (1980) Nonviral oligonucleotides at the 5' terminus of cytoplasmic influenza viral mRNA deduced from cloned complete genomic sequences. *Cell* 21 (2): 495-500.

(E)

Ehresmann, C., F. Baudin, M. Mougel, P. Romby, J. Ebel, and B. Ehresmann. (1987) Probing the structure of RNAs in solution. *NAR* 15 (22): 9109-9128.

Ehretsmann, C. P., A. J. Carpousis, and H. M. Krisch. (1992) Specificity of Escherichia coli endoribonuclease RNase E: in vivo and in vitro analysis of mutants in a bacteriophage T4 mRNA processing site. *Genes Dev* 6 (1): 149-159.

Englard, S. and S. Seifter. (1990) Precipitation Techniques. *Met Enzymol* 182: 285-300.

Esteban, R., T. Fujimura, and R. B. Wickner. (1989) Internal and terminal cis-acting sites are necessary for in vitro replication of the L-A double-stranded RNA virus of yeast. *EMBO J* 8 (3): 947-954.

(F)

Fairbanks, L.D., M. Bofill, K. Ruckemann, and H.A. Simmonds. (1995) Importance of ribonucleotide availability to proliferating T-lymphocytes from healthy humans. *J Biol Chem* 270 (15): 29682-29689.

Feigenblum, D. and R. J. Schneider. (1993) Modification of eukaryotic initiation factor 4F during infection by influenza virus. *J Virol* 67 (6): 3027-3035.

Fodor, E., D. C. Pritlove, and G. G. Brownlee. (1995) Characterization of the RNA-fork model of virion RNA in the initiation of transcription in influenza A virus. *J Virol* 69 (7): 4012-4019.

Fortes, P., A. Beloso, and J. Ortin. (1994) Influenza virus NS1 protein inhibits pre-mRNA splicing and blocks mRNA nucleocytoplasmic transport. *Embo J* 13 (3): 704-712.

Frank, K.B., J.-F. Chiou, and Y.-C. Cheng. (1984) Interaction of herpes simplex virus-induced DNA polymerase with 9-(1,3-dihydroxy-2-propoxymethyl)guanine triphosphate. *J Biol Chem* 259 (3): 1566-1569.

Fujimura, T. and R. B. Wickner. (1992) Interaction of two cis sites with the RNA replicase of the yeast L-A virus. *J Biol Chem* 267 (4): 2708-2713.

Fujiyoshi, Y., N. P. Kume, K. Sakata, and S. B. Sato. (1994) Fine structure of influenza A virus observed by electron cryo- microscopy. *Embo J* 13 (2): 318-326.

(G)

Galarza, J.M., A. Sowa, V.M. Hill, R. Skorko, and D.F. Summers. (1992) Influenza virus NP protein expressed in insect cells by a recombinant baculovirus is associated with a protein kinase activity and possesses single-stranded RNA binding activity. *Virus Res* 24: 91-106.

Gao, Y. and J. Lenard. (1995) Multimerization and transcriptional activation of the phosphoprotein (P) of vesicular stomatitis virus by casein kinase-II. *EMBO J* 14 (6): 1240-1247.

Garfinkel, M. S. and M. G. Katze. (**1993**) Translational control by influenza virus. Selective translation is mediated by sequences within the viral mRNA 5'-untranslated region. *J Biol Chem* 268 (30): 22223-22226.

Garnier, J., D.J. Osguthorpe, and B. Robson. (**1978**) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120: 97-120.

Geourjohn, C. and G. Deléage. (**1994**) SOPM: a self-optimized method for protein secondary structure prediction. *Protein Engng.* 7 (2): 157-164.

Gill, S.C. and P.H. von Hippel. (**1989**) Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 182: 319-326

Goldberg, A. L. (**1992**) The mechanism and functions of ATP-dependent proteases in bacterial and animal cells. *Eur J Biochem* 203 (1-2): 9-23.

Goody, R.S., B. Mueller, and T. Restle. (**1991**) Factors contributing to the inhibition of HIV reverse transcriptase by chain-terminating nucleotides in vitro and in vivo. *FEBS Lett* 291 (1): 1-5.

Gorbalenya, A.E. and E.V. Koonin. (**1993**) Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* 3: 419-429.

Gottesmann, S. and M.R. Maurizi. (**1992**) Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiol. Rev.* 56: 592-621.

Grosjean, H. and W. Fiers. (**1982**) Preferential codon usage in prokaryotic genes: the optimal codon- anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene* 18 (3): 199-209.

Gubareva, L.V., C.R. Penn, and R.G. Webster. (**1995**) Inhibition of replication of avian influenza viruses by the neuraminidase inhibitor 4-guanidino-2,4-dideoxy-2,3-dehydro-N-acetylneuraminic acid. *Virology* 212: 323-330.

(H)

Hagen, M., T. D. Chung, J. A. Butcher, and M. Krystal. (**1994**) Recombinant influenza virus polymerase: requirement of both 5' and 3' viral ends for endonuclease activity. *J Virol* 68 (3): 1509-1515.

Hagen, M., L. Tiley, T. D. Chung, and M. Krystal. (**1995**) The role of template-primer interactions in cleavage and initiation by the influenza virus polymerase. *J Gen Virol* 76: 603-611.

Hajnsdorf, E., O. Steier, L. Coscoy, L. Teysset, and P. Regnier. (**1994**) Roles of RNase E, RNase II and PNPase in the degradation of the rpsO transcripts of Escherichia coli: stabilizing function of RNase II and evidence for efficient degradation in an ams pnp rnb mutant. *Embo J* 13 (14): 3368-3377.

Harris, K. S., W. Xiang, L. Alexander, W. S. Lane, A. V. Paul, and E. Wimmer. (**1994**) Interaction of poliovirus polypeptide 3CDpro with the 5' and 3' termini of the poliovirus genome. Identification of viral and cellular cofactors needed for efficient binding. *J Biol Chem* 269 (43): 27004-27014.

Hay, A.J., G. Abraham, J.J. Skehel, J.C. Smith, and P. Fellner. (**1977a**) Influenza virus messenger RNAs are incomplete transcripts of the genome RNA. *Nucl. Acids Res.* 4: 4197-4209.

Hay, A. J., B. Lomniczi, A. R. Bellamy, and J. J. Skehel. (**1977b**) Transcription of the influenza virus genome. *Virology* 83 (2): 337-355.

Hay, A. J., J. J. Skehel, and J. McCauley. (**1980**) Structure and synthesis of influenza virus complementary RNAs. *Philos Trans R Soc Lond Biol* 288 (1029): 341-348.

Hay, A. J., J. J. Skehel, and J. McCauley. (**1982**) Characterization of influenza virus RNA complete transcripts. *Virology* 116 (2): 517-522.

Hay, A.J., A.J. Wolstenholme, J.J. Skehel, and M.H. Smith. (**1985**) The molecular basis of the specific anti-influenza action of amantadine. *EMBO J* 4: 3021-3024.

Hayden, F.G., J.J. Treanor, R.F. Betts, M. Lobo, J.D. Esinhart, and E.K. Hussey (**1996**) Safety and efficacy of the neuraminidase inhibitor GG167 in experimental human influenza. JAMA 275: 295-299.

Hayden, F.G. and A.J. Hay. (**1992**) Emergence and transmission of Influenza A viruses resistant to amantidine and rimantadine. *Curr. Top. Microbiol. Immunol.* 176: 119-130.

Heggeness, M.H., A. Scheid, and P.W. Choppin. (1980) Conformation of the helical nucleocapsids of paramyxoviruses and vesicular stomatitis virus: Reversible coiling and uncoiling induced by changes in salt concentration. *Proc. Natl. Acad. Sci. USA* 77 (5): 2631-2635.

Heggeness, M.H., P.R. Smith, I. Ulmanen, R.M. Krug, and P.W. Choppin. (1982) Studies on the helical nucleocapsid of Influenza virus. *Virology* 118: 466-470.

Herget, M. and C. Scholtissek. (1993) A temperature-sensitive mutation in the acidic polymerase gene of an influenza A virus alters the regulation of viral protein synthesis. *J Gen Virol* 74 (Pt 9): 1789-1794.

Herz, C., E. Stavnezer, R. Krug, and T. Gurney Jr. (1981) Influenza virus, an RNA virus, synthesizes its messenger RNA in the nucleus of infected cells. *Cell* 26 (3 Pt 1): 391-400.

Hinshaw, V. S., C. W. Olsen, N. Dybdahl-Sissoko, and D. Evans. (1994) Apoptosis: a mechanism of cell killing by influenza A and B viruses. *J Virol* 68 (6): 3667-3673.

Hirsch, M.S. and J.C. Kaplan. (1990). Antiviral agents. *Virology; edited by B.N. Fields; Raven Press, Ltd., New York*: 441-468

Holmberg, L., Y. Melander, and O. Nygard. (1992) Ribosome-bound eucaryotic elongation factor 2 protects 5S rRNA from modification. *J. Biol. Chem.* 267 (30): 21906-21910.

Honda, A., K. Mizumoto, and A. Ishihama. (1986) RNA polymerase of influenza virus. Dinucleotide-primed initiation of transcription at specific positions on viral RNA. *J Biol Chem* 261 (13): 5987-5991.

Honda, A., K. Ueda, K. Nagata, and A. Ishihama. (1987) Identification of the RNA polymerase-binding site on genome RNA of influenza virus. *J Biochem* 102 (5): 1241-1249.

Honda, A., K. Ueda, K. Nagata, and A. Ishihama. (1988) RNA polymerase of influenza virus: role of NP in RNA chain elongation. *J Biochem* 104 (6): 1021-1026.

Honda, A., J. Mukaigawa, A. Yokoiyama, A. Kato, S. Ueda, K. Nagata, M. Krystal, D. P. Nayak, and A. Ishihama. (1990) Purification and molecular structure of RNA polymerase from influenza virus A/PR8. *J Biochem* 107 (4): 624-628.

Houck, P., M. Hemphill, D. LaCroix, D. Hirsh, and N. Cox. (1995) Amantadine resistant influenza A in nursing homes. Identification of a resistant virus prior to drug use. *Arch. Intern. Med.* 155: 533-537.

Hsu, M. T., J. D. Parvin, S. Gupta, M. Krystal, and P. Palese. (1987) Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proc Natl Acad Sci U S A* 84 (22): 8140-8144.

Huang, P., D. Farquhar, and W. Plunkett. (1990a) Selective action of 3'-azido-3'-deoxythymidine- 5'-triphosphate on viral reverse transcriptases and human DNA polymerases. *J Biol Chem* 265 (20): 11914-11918.

Huang, T. S., P. Palese, and M. Krystal. (1990b) Determination of influenza virus proteins required for genome replication. *J Virol* 64 (11): 5669-5673.

(I)

Inglis, S. C., A. R. Carroll, R. A. Lamb, and B. W. Mahy. (1976) Polypeptides specified by the influenza virus genome I. Evidence for eight distinct gene products specified by fowl plague virus. *Virology* 74 (2): 489-503.

Ishihama, A., K. Mizumoto, K. Kawakami, A. Kato, and A. Honda. (1986) Proofreading function associated with the RNA-dependent RNA polymerase from influenza virus. *J Biol Chem* 261 (22): 10417-10421.

Izaurralde, E., J. Lewis, C. Gamberi, A. Jarmolowski, C. McGuigan, and I.W. Mattaj. (1995) A cap-binding protein complex mediating U snRNA export. *Nature* 376: 709-712.

(J)

Jackson, D. A., A. J. Caton, S. J. McCready, and P. R. Cook. (1982) Influenza virus RNA is synthesized at fixed sites in the nucleus. *Nature* 296 (5855): 366-368.

Jacob, G. A., J. A. Kitzmiller, and D. S. Luse. (1994) RNA polymerase II promoter strength in vitro may be reduced by defects at initiation or promoter clearance. *J Biol Chem* 269 (5): 3655-3663.

Jandrositz, A. and C. Guthrie. (**1995**) Evidence for a Prp24 binding site in U6 snRNA and in a putative intermediate in the annealing of U6 and U4 snRNAs. *EMBO J.* 14 (4): 820-832.
Jedrzejas, M.J., S. Singh, W.J. Brouillette, W.G. Laver, G.M. Air, and M. Luo. (**1995**) Structures of aromatic inhibitors of influenza virus neuraminidase. *Biochemistry* 34: 3144-3151.
Jennings, P. A., J. T. Finch, G. Winter, and J. S. Robertson. (**1983**) Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? *Cell* 34 (2): 619-627.
Jindal, H., C. Yong, G. Wilson, P. Tam, and C. Astell. (**1994**) Mutations in the NTP-binding motif of Minute Virus of Mice (MVM) NS-1 protein uncouple ATPase and DNA helicase functions. *J Biol Chem* 269 (5): 3283-3289.
Jones, I. M., P. A. Reay, and K. L. Philpott. (**1986**) Nuclear location of all three influenza polymerase proteins and a nuclear signal in polymerase PB2. *Embo J* 5 (9): 2371-2376.

(K)
Kandror, O., L. Busconi, M. Sherman, and A. L. Goldberg. (**1994**) Rapid degradation of an abnormal protein in Escherichia coli involves the chaperones GroEL and GroES. *J. Biol. Chem.* 269 (38): 23575-23582.
Kawakami, K., K. Mizumoto, and A. Ishihama. (**1983**) RNA polymerase of influenza virus. IV. Catalytic properties of the capped RNA endonuclease associated with the RNA polymerase. *Nucleic Acids Res* 11 (11): 3637-3649.
Kilbourne, E.D. (**1987**) in *Influenza; E.D. Kilbourne, ed.; Plenum, New York* : 291-332.
Kimura, N., M. Nishida, K. Nagata, A. Ishihama, K. Oda, and S. Nakada. (**1992**) Transcription of a recombinant influenza virus RNA in cells that can express the influenza virus RNA polymerase and nucleoprotein genes. *J Gen Virol* 73 (Pt 6): 1321-1328.
Kimura, N., A. Fukushima, K. Oda, and S. Nakada. (**1993**) An in vivo study of the replication origin in the influenza virus complementary RNA. *J Biochem* 113 (1): 88-92.
King, L.A. and R.D. Possee. (**1993**) The Baculovirus Expression System: A Laboratory Guide; Chapman & Hall, New York.
Kingsbury, D.W., I.M. Jones, and K.G. Murti. (**1987**) Assembly of influenza ribonucleoprotein in vitro using recombinant nucleoprotein. *Virology* 156 (2): 396-403.
Kingsbury, D.W. (**1990**) Orthomyxoviridae and their replication. *Virology, second edition; edited by B.N. Fields, D.M. Knipe et al.; Raven Press, Ltd., New York* 1: 1075-1089.
Kitts, P. A. and R. D. Possee. (**1993**) A method for producing recombinant baculovirus expression vectors at high frequency. *Biotechniques* 14 (5): 810-817.
Kobayashi, M., K. Tuchiya, K. Nagata, and A. Ishihama. (**1992**) Reconstitution of influenza virus RNA polymerase from three subunits expressed using recombinant baculovirus system. *Virus Res* 22 (3): 235-245.
Krakow, J.S., G. Rhodes, and T.M. Jovin. (**1976**). *RNA Polymerase; edited by R. Losick and M.J. Chamberlin; Cold Spring Harbor Laboratory Press, N.Y.*
Krug, R. M., M. Ueda and P. Palese (1975). Temperature-sensitive mutants of influenza WSN virus defective in virus-specific RNA synthesis. J Virol 16 (4): 790-796.
Krug, R. M., B. A. Broni, and M. Bouloy. (**1979**) Are the 5' ends of influenza viral mRNAs synthesized in vivo donated by host mRNAs? *Cell* 18 (2): 329-334.
Krug, R. M., B. A. Broni, A. J. LaFiandra, M. A. Morgan, and A. J. Shatkin. (**1980**) Priming and inhibitory activities of RNAs for the influenza viral transcriptase do not require base pairing with the virion template RNA. *Proc Natl Acad Sci U S A* 77 (10): 5874-5878.
Krug, R. M., F. V. Alonso-Caplen, I. Julkunen, and M. G. Katze. (**1989**). Expression and replication of the Influenza virus genome. In *The Influenza Viruses; edited by R.M. Krug; Plenum Press, New York*: 89-152
Krystal, M., R. Li, D. Lyles, G. Pavlakis, and P. Palese. (**1986**) Expression of the three influenza virus polymerase proteins in a single cell allows growth complementation of viral mutants. *Proc Natl Acad Sci U S A* 83 (8): 2709-2713.
Kubori, T. and N. Shimamoto. (**1996**) A branched pathway in the early stage of transcription by Escherichia coli RNA polymerase. *J. Mol. Biol.* 256 (3): 449-457.

(L)
Laemmli, U.K. (**1970**) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature* 227: 680-685.
Lain, S., J. L. Riechmann, and J. A. Garcia. (**1990**) RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Res* 18 (23): 7003-7006.
Lamb, R.A. and P.W. Choppin. (**1977**) Synthesis of Influenza virus polypeptides in cells resistant to alpha-amanitin: evidence for the involvement of cellular RNA polymerase II in virus replication. *J. Virol.* 23 (3): 816-819.
Lamb, R.A., C.-J. Lai, and P.W. Choppin. (**1981**) Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. *Proc. Natl. Acad. Sci. USA* 77: 4170-4174.
Lamb, R.A. and P.W. Choppin. (**1983**) The gene structure and replication of influenza virus. *Ann. Rev. Biochem.* 52: 467-506.
Lamb, R.A. (**1989**) Genes and proteins of the Influenza viruses. *The Influenza Viruses Krug, R. M. ed.; Plenum Press, New York* : 1-87.
Lazarowitz, S. G., R. W. Compans, and P. W. Choppin. (**1971**) Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. *Virology* 46 (3): 830-843.
Lee, T. G., J. Tomita, A. G. Hovanessian, and M. G. Katze. (**1990**) Purification and partial characterization of a cellular inhibitor of the interferon-induced protein kinase of Mr 68,000 from influenza virus- infected cells. *Proc Natl Acad Sci U S A* 87 (16): 6208-6212.
Lee, T. G., J. Tomita, A. G. Hovanessian, and M. G. Katze. (**1992**) Characterization and regulation of the 58,000-dalton cellular inhibitor of the interferon-induced, dsRNA-activated protein kinase. *J Biol Chem* 267 (20): 14238-14243.
Li, H., S. Dalal, J. Kohler, J. Vilardell, and S. A. White. (**1995**) Characterization of the pre-mRNA binding site for yeast ribosomal protein L32: the importance of a purine-rich internal loop. *J Mol Biol* 250 (4): 447-459.
Li, X. and P. Palese. (**1994**) Characterization of the polyadenylation signal of influenza virus RNA. *J Virol* 68 (2): 1245-1249.
Licheng, S., D.F. Summers, Q. Peng, and J.M. Galarza. (**1995**) Influenza A virus polymerase subunit PB2 is the endonuclease which cleaves host cell mRNA and functions only as the trimeric enzyme. *Virology* 208: 38-47.
Lopez-Turiso, J. A., C. Martinez, T. Tanaka, and J. Ortin. (**1990**) The synthesis of influenza virus negative-strand RNA takes place in insoluble complexes present in the nuclear matrix fraction. *Virus Res* 16 (3): 325-337.
Lu, Y., X. Y. Qian, and R. M. Krug. (**1994**) The influenza virus NS1 protein: a novel inhibitor of pre-mRNA splicing. *Genes Dev* 8 (15): 1817-1828.
Lu, Y., M. Wambach, M. G. Katze, and R. M. Krug. (**1995**) Binding of the influenza virus NS1 protein to double-stranded RNA inhibits the activation of the protein kinase that phosphorylates the eIF-2 translation initiation factor. *Virology* 214 (1): 222-228.
Luo, G. X., W. Luytjes, M. Enami, and P. Palese. (**1991**) The polyadenylation signal of influenza virus RNA involves a stretch of uridines followed by the RNA duplex of the panhandle structure. *J Virol* 65 (6): 2861-2867.
Luytjes, W., M. Krystal, M. Enami, J. D. Pavin, and P. Palese. (**1989**) Amplification, expression, and packaging of foreign gene by influenza virus. *Cell* 59 (6): 1107-1113.

(M)
Mackie, G.A. and J.L. Genereaux. (**1993**) The role of RNA structure in determining RNase E-dependent cleavage sites in the mRNA for ribosomal protein S20 in vitro. *J. Mol. Biol.* 234: 998-1012.
Mar, E.-C., J.-F. Chiou, Y.-C. Cheng, and E.-S. Huang. (**1985**) Inhibition of cellular DNA polymerase a and human cytomegalovirus-induced DNA polymerase by the triphosphates of 9-(2-hydroxyethoxymethyl)guanine and 9-(1,3-dihydroxy-2-propoxymethyl)guanine. *J Virol* 53 (3): 776-780.
Marsh, M. and A. Helenius. (**1989**) Virus entry into animal cells. *Adv. Virus Res.* 36: 107-151.

Martin, K. and A. Helenius. (**1991a**) Nuclear transport of influenza virus ribonucleoproteins: the viral matrix protein (M1) promotes export and inhibits import. *Cell* 67 (1): 117-130.
Martin, K. and A. Helenius. (**1991b**) Transport of incoming influenza virus nucleocapsids into the nucleus. *J Virol* 65 (1): 232-244.
Matsuura, Y., R.D. Possee, H.A. Overton, and D.H.L. Bishop. (**1987**) Baculovirus expression vectors: the requirements for high level expression of proteins, including glycoproteins. *J Gen Virol* 68: 1233-1250
Mattes, R. (**1993**) Principles of gene expression. *in H.-J. Rehm and G. Reed (Eds.), Biotechnology, Vol. II, VCH, Weinheim* : 233-256.
Matunis, E.L., M.J. Matunis, and G. Dreyfuss. (**1993**) Association of individual hnRNP proteins and snRNPs with nascent transcripts. *J. Cell Biol.* 121: 219-228.
McCormick, J.B., I.J. King, P.A. Webb, and et al. (**1986**) Lassa fever: effective therapy with ribavirin. *N. Engl. J. Med.* 314: 20-26.
McGeoch, D. and N. Kitron. (**1975**) Influenza virion RNA-dependent RNA polymerase: stimulation by guanosine and related compounds. *J. Virol.* 15 (4): 686-695.
McIntosh, K., S.G. Kurachek, M.M. Cairns, J.G. Burns, and B. Goodspeed. (**1984**) Treatment of respiratory viral infection in an immunodeficient infant with ribavirin aerosol. *Am J Dis Child* 138: 305-308.
Mena, I., S. de la Luna, C. Albo, J. Martin, A. Nieto, J. Ortin, and A. Portela. (**1994**) Synthesis of biologically active influenza virus core proteins using a vaccinia virus-T7 RNA polymerase expression system. *J Gen Virol* 75: 2109-2114.
Michael, W.M. and G. Dreyfuss. (**1996**) Distinct domains in ribosomal protein L5 mediate 5S rRNA binding and nucleolar localization. *J. Biol. Chem.* 271 (19): 11571-11574.
Moazed, D., J.M. Robertson, and H.F. Noller. (**1988**) Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S RNA. *Nature* 334: 362-364.
Murphy, B.R. and R.G. Webster. (**1990**) Orthomyxoviruses. *Virology, second edition; edited by B.N. Fields, D.M. Knipe et al.; Raven Press, Ltd., New York* 1: 1091-1152.
Murti, K. G., R. G. Webster, and I. M. Jones. (**1988**) Localization of RNA polymerases on influenza viral ribonucleoproteins by immunogold labeling. *Virology* 164 (2): 562-566.

(N)
Nakagawa, Y., N. Kimura, T. Toyoda, K. Mizumoto, A. Ishihama, K. Oda, and S. Nakada. (**1995**) The RNA polymerase PB2 subunit is not required for replication of the influenza virus genome but is involved in capped mRNA synthesis. *J Virol* 69 (2): 728-733.
Nakagawa, Y., K. Oda, and S. Nakada. (**1996**) The PB1 subunit alone can catalyze cRNA synthesis, and the PA subunit in addition to the PB1 subunit is required for viral RNA synthesis in replication of the influenza virus genome. *J Virol* 70 (9): 6390-6394.
Nieto, A., S. de la Luna, J. Barcena, A. Portela, J. Valcarcel, J. A. Melero, and J. Ortin. (**1992**) Nuclear transport of influenza virus polymerase PA protein. *Virus Res* 24 (1): 65-75.

(O)
O'Reilly, D.R., L.K. Miller, and V.A. Luckow. (**1992**) Baculovirus Expression Vectors: Á Laboratory Manual. *W.H. Freeman and Company* .
Okazaki, K., Y. Kawaoka, and R.G. Webster. (**1989**) Evolutionary pathways of the PA genes of Influenza A viruses. *Virology* 172: 601-608.
Osumi-Davis, P.A., M.C. de Aguilera, R.W. Woody, and A. Young Moon Woody. (**1992**) Asp537, Asp812 are essential and Lys631, His811 are catalytically significant in bacteriophage T7 RNA polymerase activity. *J Mol Biol* 226: 37-45.
Oubridge, C., N. Ito, P.R. Evans, C. Teo, and K. Nagai. (**1994**) Crystal structure at 1.92A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* 372: 432-438.

(P)
Pace, C.N., F. Vajdos, L. Fee, G. Grimsley, and T. Gray. (**1995**) How to measure and predict the molar absorption coefficient of a protein. *Protein Science* 4: 2411-2423.

Palese, P., K. Tobita, and M. Ueda. (1974) Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. *Virology* 61: 397-410.
Palese, P., M.B. Ritchey and J.L. Schulman (1977). P1 and P3 proteins of influenza virus are required for complementary RNA synthesis. J Virol 21 (3): 1187-1195.
Parker, W.B., E.L. White, S.C. Shaddix, L.J. Ross, R.W. Buckheit Jr., J.M. Germany, J.A. Secrist III, R. Vince, and W.M. Shannon. (1991) Mechanism of inhibition of human immunodeficiency virus type 1 reverse transcriptase and human DNA polymerases α, β and γ by the 5'-triphosphates of carbovir, 3'-azido-3'-deoxythymidine, 2',3'-dideoxyguanosine and 3'-deoxythymidine. *J Biol Chem* 266 (3): 1754-1762.
Pata, J.D., S.C. Schultz, and K. Kirkegaard. (1995) Functional oligomerization of poliovirus RNA-dependent RNA polymerase. *RNA* 1: 466-477.
Peattie, D. A. and W. Gilbert. (1980) Chemical probes for higher-order structure in RNA. *Proc Natl Acad Sci U S A* 77 (8): 4679-4682.
Perrault, J. and P.W. McLear. (1984) ATP dependence of vesicular stomatitis virus transcription initiation and modulation by mutation in the nucleocapsid protein. *J Virol* 51 (3): 635-642.
Piñol-Roma, S. and G. Dreyfuss. (1992) Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* 355: 730-732.
Pinto, L.H., L.J. Holsinger, and R.A. Lamb. (1992) Influenza virus M2 protein has ion channel activity. *Cell* 69: 517-528.
Plotch, S.J. and R.M. Krug. (1977) Influenza virion transcriptase: synthesis in vitro of large, polyadenylic acid-containing complementary RNA. *J Virol* 21 (1): 24-34.
Plotch, S.J. and R.M. Krug. (1978) Segments of influenza virus complementary RNA synthesized in vitro. *J. Virol.* 25 (2): 579-586.
Plotch, S.J., M. Bouloy, and R.M. Krug. (1979) Transfer of 5'-terminal cap of globin mRNA to influenza viral complementary RNA during transcription in vitro. *Proc. Natl. Acad. Sci. USA* 76 (4): 1618-1622.
Plotch, S.J., M. Bouloy, I. Ulmanen, and R.M. Krug. (1981) A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* 23 (3): 847-858.
Poch, O., I. Sauvaget, M. Delarue, and N. Tordo. (1990) Identification of four conserved motifs among the RNA-dependent polymerases encoding elements. *EMBO J.* 8: 3867-3874.
Pons, M.W., I.T. Schulze, G.K. Hirst, and R. Hauser. (1969) Isolation and characterization of the ribonucleoprotein of influenza virus. *Virology* 39 (2): 250-259.

(Q)
Qiu, Y. and R. M. Krug. (1994) The influenza virus NS1 protein is a poly(A)-binding protein that inhibits nuclear export of mRNAs containing poly(A). *J. Virol.* 68 (4): 2425-2432.
Qiu, Y., M. Nemeroff, and R. M. Krug. (1995) The influenza virus NS1 protein binds to a specific region in human U6 snRNA and inhibits U6-U2 and U6-U4 snRNA interactions during splicing. *Rna* 1 (3): 304-316.
Quadt, R., M. Ishikawa, M. Janda, and P. Ahlquist. (1995) Formation of brome mosaic virus RNA-dependent RNA polymerase in yeast requires coexpression of viral proteins and viral RNA. *Proc. Natl. Acad. Sci. U S A* 9295: 4892-4896.

(R)
Rao, L., M. Debbas, P. Sabbatini, D. Hockenbery, S. Korsmeyer, and E. White. (1992) The adenovirus E1A proteins induce apoptosis, which is inhibited by the E1B 19-kDa and Bcl-2 proteins. *Proc Natl Acad Sci U S A* 89 (16): 7742-7746.
Robertson, J. S. (1979) 5' and 3' terminal nucleotide sequences of the RNA genome segments of influenza virus. *Nucleic Acids Res* 6 (12): 3745-3757.
Robertson, J.S., M. Schubert, and R.A. Lazzarini. (1981) Polyadenylation sites for influenza virus mRNA. *J. Virol.* 38: 157-163.
Rochovansky, O. (1976) RNA synthesis by ribonucleoprotein-polymerase complexes isolated from influenza virus. *Virology* 73: 327-338.
Rost, B. and C. Sander. (1994) Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins* 19: 55-72.

Rott, R., S. Saber, and C. Scholtissek. (1965) Effect on myxovirus of mitomycin C, actinomycin D and pretreatment of the host cell with ultraviolet light. *Nature* 205: 1187-1190.
Rott, R. and C. Scholtissek. (1970) Specific inhibition of influenza replication by alpha-amanitin. *Nature* 228: 56.
Ruben, F.L. (1987) Prevention and control of Influenza: Role of vaccine. *Am. J. Med.* 82 (Suppl.6A): 31-34.
Ruigrok, R.W.H., P.J. Andree, R.A.M. Hooft van Huysduynen, and J.E. Mellema. (1984) Characterization of three highly purified influenza virus strains by electron microscopy. *J Gen Virol* 65: 799-802.
Ruigrok, R. W., L. J. Calder, and S. A. Wharton. (1989) Electron microscopy of the influenza virus submembranal structure. *Virology* 173 (1): 311-316.
Ruigrok, R.W.H. and F. Baudin. (1995) Structure of Influenza virus ribonucleoprotein particles.II. Purified RNA-free Influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact Influenza virus ribonucleoprotein particles. *J. Gen. Virol.* 76: 1009-1014.

(S)
Sakaguchi, T., G.P. Leser, and R.A. Lamb. (1996) The ion channel activity of the Influenza virus M2 protein affects transport through the Golgi apparatus. *J. Cell Biol.* 133 (4): 733-747.
Sambrook, J., E.F. Fritsch, and T. Maniatis. (1989) *Molecular Cloning: A Laboratory Manual; edited by C. Nolan; Cold Spring Harbour Laboratory Press, Cold Spring Harbour.*
Sanz-Ezquerro, J. J., S. de la Luna, J. Ortin, and A. Nieto. (1995) Individual expression of influenza virus PA protein induces degradation of coexpressed proteins. *J Virol* 6964 (42034): 2420-2426.
Sanz-Ezquerro, J.J., T. Zuercher, S. de la Luna, J. Ortin, and A. Nieto. (1996) The amino-terminal one third of the Influenza virus PA protein is responsible for the induction of proteolysis. *J. Virol.* 70 (3): 1905-1911.
Schenk, P.M., S. Baumann, R. Mattes, and H. Steinbiss. (1995) Improved high-level expression system for eucaryotic genes in Escherichia coli using T7 RNA polymerase and rare Arg-tRNAs. *BioTechniques* 19: 196-200.
Scherly, D., W. Boelens, W.J. van Venrooij, N.A. Dathan, J. Hamm, and I.W. Mattaj. (1989) Identification of the RNA binding segment of human U1A protein and definition of its binding site on U1 snRNA. *EMBO J.* 8 (13): 4163-4170.
Scherly, D., W. Boelens, N.A. Dathan, W.J. van Venrooij, and I.W. Mattaj. (1990) Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B" and their cognate RNAs. *Nature* 345: 502-506.
Schulman, J.L. and P. Palese. (1975) Susceptibility of different strains of influenza A virus to the inhibitory effects of 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA). *Virology* 63: 98-104.
Schuppli, D., I. Barrera, and H. Weber. (1994) Identification of recognition elements on bacteriophage Q beta minus strand RNA that are essential for template activity with Q beta replicase. *J Mol Biol* 243 (5): 811-815.
Seong, B. L. and G. G. Brownlee. (1992) A new method for reconstituting influenza polymerase and RNA in vitro: a study of the promoter elements for cRNA and vRNA synthesis in vitro and viral rescue in vivo. *Virology* 186 (1): 247-260.
Severini, A., X.Y. Liu, J.S. Wilson, and D.L. Tyrrell. (1995) Mechanism of inhibition of duck hepatitis B virus polymerase by (-)-beta-L-2',3'-dideoxy-3'-thiacytidine. *Antimicrob Agents Chemother* 39: 1430-1435.
Shapiro, G. I., T. Gurney Jr., and R. M. Krug. (1987) Influenza virus gene expression: control mechanisms at early and late times of infection and nuclear-cytoplasmic transport of virus-specific RNAs. *J Virol* 61 (3): 764-773.
Shapiro, G. I. and R. M. Krug. (1988) Influenza virus RNA replication in vitro: synthesis of viral template RNAs and virion RNAs in the absence of an added primer. *J Virol* 62 (7): 2285-2290.
Sharp, P.M. and E. Cowe. (1991) Synonymous codon usage in Saccharomyces cerevisiae. *Yeast* 7: 657-678.

Shi, L., D.F. Summers, Q. Peng, and J.M. Galarza. (1995) Influenza A virus RNA polymerase subunit PB2 is the endonuclease which cleaves host cell mRNA and functions only as the trimeric enzyme. *Virology* 208: 38-47.

Skehel, J. J. and G. C. Schild. (1971) The polypeptide composition of influenza A viruses. *Virology* 44 (2): 396-408.

Skehel, J. J. (1972) Polypeptide synthesis in influenza virus-infected cells. *Virology* 49 (1): 23-36.

Skehel, J. J. and A. J. Hay. (1978) Nucleotide sequences at the 5' termini of influenza virus RNAs and their transcripts. *Nucleic Acids Res* 5 (4): 1207-1219.

Skehel, J.J., P.M. Bayley, E.B. Brown, S.R. Martin, M.D. Waterfield, J.M. White, I.A. Wilson, and D.C. Wiley. (1982) Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proc Natl Acad Sci USA* 79: 968-972.

Smith, F.I. and P. Palese. (1989) Variation in Influenza virus genes. *The Influenza Viruses Krug, R. M. ed.; Plenum Press, New York* : 319-359.

Sonenberg, N. (1993) Translation factors as effectors of cell growth and tumorigenesis. *Curr. Op. Cell Biol.* 5: 955-960.

Sousa, R., D. Patra, and E. M. Lafer. (1992) Model for the mechanism of bacteriophage T7 RNAP transcription initiation and termination. *J Mol Biol* 224 (2): 319-334.

Spanjaard, R.A., K. Chen, J.R. Walter, and J. van Duin. (1990) Frameshift suppression at tandem AGA and AGG codons by cloned tRNA genes: assigning a codon to argU tRNA and T4 tRNA-Arg. *Nucleic Acids Res.* 18: 5031-5036.

St. Angelo, C., G. E. Smith, M. D. Summers, and R. M. Krug. (1987) Two of the three influenza viral polymerase proteins expressed by using baculovirus vectors form a complex in insect cells. *J. Virol.* 61 (2): 361-365.

St.Clair, M.H., C.A. Richards, T. Spector, K.J. Weinhold, W.H. Miller, A.J. Langlois, and P.A. Furman. (1987) 3'-azido-3'-deoxythymidine triphosphate as an inhibitor and substrate of purified human immunodeficiency virus reverse transcriptase. *Antimicrob Agents Chemother* 31: 1972-1977.

Stoeckle, M. Y., M. W. Shaw, and P. W. Choppin. (1987) Segment-specific and common nucleotide sequences in the noncoding regions of influenza B virus genome RNAs. *Proc Natl Acad Sci U S A* 84 (9): 2703-2707.

Stranden, A. M., P. Staeheli, and J. Pavlovic. (1993) Function of the mouse Mx1 protein is inhibited by overexpression of the PB2 protein of influenza virus. *Virology* 197 (2): 642-51.

Stridh, S. (1983) Determination of ribonucleoside triphosphate pools in influenza A virus infected MDCK cells. *Arch Virol* 77: 223-229.

Stridh, S. and R. Datema. (1984) Mode of interference of trisodium phosphonoformate (INN: foscarnet sodium), with influenza virus mRNA synthesis. *Virology* 135 (1): 293-296.

Szewczyk, B., W. G. Laver, and D. F. Summers. (1988) Purification, thioredoxin renaturation, and reconstituted activity of the three subunits of the influenza A virus RNA polymerase. *Proc Natl Acad Sci U S A* 85 (21): 7907-7911.

(T)

Takizawa, T., S. Matsukawa, Y. Higuchi, S. Nakamura, Y. Nakanishi, and R. Fukuda. (1993) Induction of programmed cell death (apoptosis) by influenza virus infection in tissue culture cells. *J Gen Virol* 74 (Pt 11): 2347-2355.

Takizawa, T., R. Fukuda, T. Miyawaki, K. Ohashi, and Y. Nakanishi. (1995) Activation of the apoptotic Fas antigen-encoding gene upon influenza virus infection involving spontaneously produced beta-interferon. *Virology* 209 (2): 288-296.

Tashiro, M., P. Ciborowski, M. Reinacher, G. Pulverer, H.-D. Klenk, and R. Rott. (1987) Synergistic role of staphylococcal proteases in the induction of Influenza virus pathogenicity. *Virology* 157: 421-430.

Testa, D. and A.K. Banerjee. (1979) Initiation of RNA synthesis in vitro by vesicular stomatitis virus. *J Biol Chem* 254 (6): 2053-2058.

Tiley, L. S., M. Hagen, J. T. Matthews, and M. Krystal. (1994) Sequence-specific binding of the influenza virus RNA polymerase to sequences located at the 5' ends of the viral RNAs. *J Virol* 68 (8): 5108-5116.

Tisdale, M., G. Appleyard, J. Tuttle, D. Nelson, S. Nusinoff-Lehrmann, W. Al Nakib, J. Stables, D. Purifoy, K. Powell, and G. Darby. (1993) Inhibition of influenza A and B viruses by 2'-deoxy-2'-fluororibosides. *Antiviral Chemistry &Chemotherapy* 4: 281-287.

Tisdale, M., M. Ellis, K. Klumpp, S. Court, and M. Ford. (1995) Inhibition of Influenza virus transcription by 2'-Deoxy-2'-Fluoroguanosine. *Antimicrobial Agents and Chemotherapy* 39 (11): 2454-2458.

Tomassini, J., H. Selnick, M. E. Davies, M. E. Armstrong, J. Baldwin, M. Bourgeois, J. Hastings, D. Hazuda, J. Lewis, W. McClements, and al et. (1994) Inhibition of cap (m7GpppXm)-dependent endonuclease of influenza virus by 4-substituted 2,4-dioxobutanoic acid compounds. *Antimicrob Agents Chemother* 38 (12): 2827-2837.

(U)

Ulmanen, I., B.A. Broni, and R.M. Krug. (1981) Role of two of the influenza virus core P proteins in recognizing cap 1 structures (m7GpppNm) on RNAs and in initiating viral RNA transcription. *Proc. Natl. Acad. Sci. USA* 78: 7355-7359.

Ulmanen, I., B.A. Broni, and R.M. Krug. (1983) Influenza virus temperature-sensitive cap (m7GpppNm)-dependent endonuclease. *J. Virol.* 45: 27-35.

(V)

von Itzstein, M., J.G. Barry, and A.K.J. Chong. (1993a) The development of potential anti-influenza drugs. *Curr. Opin. Thera. Patents* 3 (12): 1755-1762.

von Itzstein, M., W. Y. Wu, G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. Van Phan, M. L. Smythe, H. F. White, S. W. Oliver, et. al. (1993b) Rational design of potent sialidase-based inhibitors of influenza virus replication [see comments]. *Nature* 363 (6428): 418-423.

(W)

Walker, J.E., M. Saraste, M.J. Runswick, and N.J. Gay. (1982) Distantly related sequences in the α- and β-subunits of ATP synthetase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1: 945-951.

Warrener, P. and M. S. Collett. (1995) Pestivirus NS3 (p80) protein possesses RNA helicase activity. *J. Virol.* 69 (3): 1720-1726.

Wennborg, A., B. Sohlberg, D. Angerer, G. Klein, and A. von Gabain. (1995) A human RNase E-like activity that cleaves RNA sequences involved in mRNA stability control. *Proc Natl Acad Sci U S A* 92 (16): 7322-7326.

White, D.O. (1974) *Curr. Top. Microbiol. Immunol.* 63: 1-48.

Whittaker, G., M. Bui, and A. Helenius. (1996) Nuclear trafficking of Influenza virus ribonucleoproteins in heterokaryons. *J. Virol.* 70 (5): 2743-2756.

Wohlrab, F., A.T. Jamieson, J. Hay, R. Mengel, and W. Guschlbauer. (1985) The effect of 2'-fluoro-2'-deoxycytidine on herpes virus growth. *Biochim Biophys Acta* 824: 233-242.

(Y)

Yamashita, M., M. Krystal, and P. Palese. (1989) Comparison of the three large polymerase proteins of influenza A, B, and C viruses. *Virology* 171 (2): 458-466.

Yoshida, M. and T. Amano. (1995) A common topology of proteins catalyzing ATP triggered reactions. *FEBS Letters* 359: 1-5.

Young, R.J. and J. Content. (1971) 5' terminus of the Influenza virus RNA. *Nature* 230: 140-142.

(Z)

Zebedee, S.L. and R.A. Lamb. (1988) Influenza A virus M2 protein: monoclonal antibody restriction of virus growth and detection of M2 in virions. *J Virol* 62 (8): 2762-2772.

Zhirnov, O. P. (1990) Solubilization of matrix protein M1/M from virions occurs at different pH for orthomyxo- and paramyxoviruses. *Virology* 176 (1): 274-279.